Global Biogeochemical Cycles

RESEARCH ARTICLE
10.1029/2024GB008127

High-Resolution Variability of the Ocean Carbon Sink

Luke Gregor[1,2], **Jamie Shutler**[3], **and Nicolas Gruber**[1]

[1]Environmental Physics, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zürich, Switzerland, [2]Swiss Data Science Center, ETH Zürich and EPFL, Zürich, Switzerland, [3]Centre for Geography and Environmental Science, University of Exeter, Cornwall, UK

Key Points:

- Surface ocean concentration and sea-air fluxes of carbon dioxide are estimated at 8-day, quarter degree resolution with a neural network
- Variability at subseasonal timescales contributes substantially to the total variability of the ocean carbon sink
- The high-resolution data provide novel observational insights into regional and ephemeral processes, such as upwelling and hurricanes

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. Gregor,
luke.gregor@sdsc.ethz.ch

Citation:

Gregor, L., Shutler, J., & Gruber, N. (2024). High-resolution variability of the ocean carbon sink. *Global Biogeochemical Cycles*, *38*, e2024GB008127. https://doi.org/10.1029/2024GB008127

Received 1 FEB 2024
Accepted 19 JUL 2024

Author Contributions:

Conceptualization: Luke Gregor, Nicolas Gruber
Data curation: Luke Gregor
Formal analysis: Luke Gregor
Funding acquisition: Jamie Shutler, Nicolas Gruber
Investigation: Luke Gregor, Jamie Shutler, Nicolas Gruber
Methodology: Luke Gregor
Project administration: Jamie Shutler, Nicolas Gruber
Resources: Luke Gregor, Nicolas Gruber
Software: Luke Gregor
Supervision: Nicolas Gruber
Validation: Luke Gregor
Visualization: Luke Gregor
Writing – original draft: Luke Gregor

Abstract Measurements of the surface ocean fugacity of carbon dioxide (fCO_2) provide an important constraint on the global ocean carbon sink, yet the gap-filling products developed so far to cope with the sparse observations are relatively coarse (1° × 1° by 1 month). Here, we overcome this limitation by using a novel combination of machine learning-based methods and target transformations to estimate surface ocean fCO_2 and the associated sea-air CO_2 fluxes (FCO_2) globally at a resolution of 8-day by 0.25° × 0.25° (8D) over the period 1982 through 2022. Globally, the method reconstructs fCO_2 with accuracy similar to that of low-resolution methods (~19 µatm), but improves it in the coastal ocean. Although global ocean CO_2 uptake differs little, the 8D product captures 15% more variance in FCO_2. Most of this increase comes from the better-represented subseasonal scale variability, which is largely driven by the better-resolved variability of the winds, but also contributed to by the better-resolved fCO_2. The high-resolution fCO_2 is also capable of capturing the signal of short-lived regional events such as hurricanes. For example, the 8D product reveals that fCO_2 was at least 25 µatm lower in the wake of Hurricane Maria (2017), the result of a complex interplay between the decrease in temperature, the entrainment of carbon-rich waters, and an increase in primary production. By providing new insights into the role of higher frequency variations of the ocean carbon sink and the underlying processes, the 8D product fills an important gap.

Plain Language Summary The ocean is important for the climate, as it takes up about a quarter of the carbon dioxide (CO_2) we release into the atmosphere every year. To determine this carbon sink, we measure the levels of carbon dioxide at the surface of the ocean. However, these measurements are limited to the locations where ships and other platforms measure CO_2, leaving gaps in our understanding. To fill in these gaps, statistical methods are used, but previous approaches lack fine-scale detail. We overcome this limitation with a neural network approach that estimates CO_2 in more detail, with estimates every 8 days, at 25 km, compared to the previous 100 km monthly estimates. Globally, our method is as accurate as the previous methods, being slightly more accurate in coastal areas. Although the total amount of carbon the ocean absorbs globally remains consistent, our results show more variability. Our method also detects short-lived local events, such as hurricanes. For example, after Hurricane Maria in 2017, carbon dioxide concentrations at the ocean surface were substantially lower. Overall, our detailed results give us new information on the small-scale changes in the ocean carbon sink and help us fill a gap in our understanding of surface ocean CO_2.

1. Introduction

The global ocean plays a pivotal role in limiting global warming by having absorbed approximately 25% of the anthropogenic carbon dioxide (CO_2) emissions over the past two centuries (Gruber, Clement, et al., 2019; Khatiwala et al., 2013; Müller et al., 2023; Sabine et al., 2004). Although this uptake fraction has remained remarkably stable over time (Friedlingstein et al., 2022), the magnitude of the ocean carbon sink has varied substantially around this trend (Bennington et al., 2022; DeVries et al., 2023; Gruber, Landschützer, & Lovenduski, 2019; Landschützer et al., 2015, 2016; McKinley et al., 2020; Rödenbeck et al., 2022). The strongest evidence supporting this variability comes from observations of the surface ocean CO_2 concentration (generally expressed in terms of its fugacity, fCO_2), from which the sea-air CO_2 flux (FCO_2) can be inferred (Fay et al., 2021; R. H. Wanninkhof, 2014). Since the fCO_2 observations are sparse in time and space, gap-filling techniques are required to map them to the time-and-space continuous product needed for estimating the strength of the global ocean carbon sink over time (Fay et al., 2021; Rödenbeck et al., 2015). Most often, statistical or machine learning techniques are used to fill this gap (Chau et al., 2022; Gloege et al., 2022; Gregor et al., 2019; Iida et al., 2021; Landschützer et al., 2013; Rödenbeck et al., 2015; Telszewski et al., 2009), although geospatial and data assimilation-type methods are also used (Bennington et al., 2022; Rödenbeck et al., 2022). Owing to their

Writing – review & editing: Luke Gregor, Jamie Shutler, Nicolas Gruber

global nature and good temporal coverage, satellite observations have proven to be key enablers for all gap-filling methods (Shutler et al., 2020, 2024).

Typically, these fCO_2 gap-filled products are produced at a monthly 1° × 1° resolution (henceforth denoted by 1M). This resolution has proven to be sufficient to constrain the global ocean uptake of CO_2 and its trends and variations (DeVries et al., 2023; Gloege et al., 2021; Gruber et al., 2023; Landschützer et al., 2016) as well as the seasonal variations and their changes over time (Landschützer et al., 2018; Rodgers et al., 2023). As a result, these products have played an important role in global assessments, especially those of the Global Carbon Budget (GCB) (Friedlingstein et al., 2022, 2023; Hauck et al., 2023). But there is a large amount of variability that current fCO_2-products cannot capture, especially at the regional and subseasonal scale. Modeling studies and observations at these scales have regularly revealed fCO_2 variations that exceed those seen in the 1M products (Arruda et al., 2015; Friederich et al., 2008; Nicholson et al., 2022; Resplandy et al., 2024; Turi et al., 2014; Yu et al., 2020). These variations are driven by finer-scale temporal features such as storms and upwelling events, but also by finer-scale spatial features, such as those associated with mesoscale circulation or strong fronts. In addition, high-frequency variations in surface winds contribute to high frequency variations in the sea-air CO_2 fluxes as well (Whitt et al., 2019).

Our current ability to constrain such high-resolution FCO_2 and fCO_2 variability from observations is very limited. This is, in no small part, a consequence of the segmented way the ocean CO_2 system is currently sampled (R. Wanninkhof et al., 2019). On the one hand, we have underway CO_2 measurements from ships that give a good perspective of the variability in space along a limited number of survey lines (Jones et al., 2012). On the other hand, we have time series observations from a few sites that provide detailed temporal information (Bates et al., 2014; Ye et al., 2020). But we rarely have observations that cover both time and space in a synoptic manner.

The few observation-based studies clearly point to the scale of the challenge. Regarding high-frequency temporal variability, a glider-based study in the Southern Ocean found that mid-latitude cyclones can induce variability up to 20 µatm within a frequency range of 1–10 days (Nicholson et al., 2022). Another glider-based study suggested that pCO_2 measurements need to be taken every 3 days to constrain uncertainty in dynamically variable regions (Monteiro et al., 2015). Beyond the Southern Ocean, intense sporadic events like tropical cyclones and hurricanes can cause pCO_2 fluctuations as large as 50 µatm within a two-day window (Bates et al., 1998; Koch et al., 2009; Yu et al., 2020). Mooring-based studies from various locations have also revealed variations of more than 50 µatm within days to weeks (Leinweber et al., 2009; Pardo et al., 2019; Sutton et al., 2014, 2017; Torres et al., 2021).

Regarding high-resolution spatial variability, gradients exceeding several tens of µatm over tens of kilometers are regularly encountered along ship-tracks, especially in dynamic regions such as the Southern Ocean and boundary current regions. This leads to short spatial autocorrelation length scales (Jones et al., 2012). Although the length scales for FCO_2 and fCO_2 are typically around 100 km or more in open ocean gyre regions, they decrease to less than 50 km in these dynamic regions (Jones et al., 2012; Murphy et al., 2001). These findings are corroborated by studies using drifters and uncrewed surface vehicles in the Northeastern Atlantic and the Southern Ocean, which have reported spatial gradients of pCO_2 on the order of 10 µatm over 20 km (Boutin et al., 2008; Merlivat et al., 2009; Sutton et al., 2021).

Better resolution of the fine-scale variations of the sea-air CO_2 fluxes matters for multiple reasons. First, it allows us to better resolve a number of key processes that govern the ocean uptake of CO_2, providing novel insights into how the ocean carbon sink functions. Second, it allows us to assess the role of potential aliasing effects resulting from missing variability, potentially aliasing our estimate of global ocean carbon uptake (Koch et al., 2009). Third, such high-resolution products can also provide critical constraints for assessing the impact of natural or man-made perturbations, such as those associated with marine heatwaves (Mignot et al., 2022) or the purposeful release of alkaline substances to enhance the oceanic uptake of atmospheric CO_2 (González & Ilyina, 2016; Lenton et al., 2018).

First attempts to cover global finer-scale variability than 1M in gap-filled fCO_2 products were undertaken by Rödenbeck et al. (2014) for the CarboScope Mixed-Layer Scheme (CarboScope-MLS) and by Chau et al. (2024) for the CMEMS-FFNN product. The CarboScope-MLS, while having higher temporal resolution (daily), suffers from its coarse spatial resolution of >2°. CMEMS-FFNN offers superior spatial resolution, but resolves fCO_2 only at monthly resolution. These limitations are, in part, technological in nature. For example, the commonly used Surface Ocean CO_2 Atlas (SOCAT) provides a monthly 1° × 1° resolution gridded product alongside the

ungridded cruise tracks (Sabine et al., 2013). Similarly, remote sensing and reanalysis products are often available at monthly resolutions in addition to daily files. However, there are some studies that have estimated high-resolution fCO_2 (~5 km) at regional scales, for example, the Gulf of Mexico (Chen et al., 2019) and the South China Sea (Sharp et al., 2022; Song et al., 2023). The regional constraint of these studies allows for simpler machine learning architecture (e.g., no clustering) that is able to predict fCO_2 with greater fidelity than global approaches.

In this study, we aim to bridge the "high-resolution gap" in current fCO_2 products by generating estimates at an 8-day, 0.25° × 0.25° (henceforth referred to as 8D) resolution for both fCO_2 and FCO_2. To achieve this, we introduce a novel combination of machine learning-based methods and also use a number of target transformations. This allows us to improve the resolution of our previously published OceanSODA-ETHZ product (Gregor & Gruber, 2021), which had a resolution of 1° × 1° by monthly to the desired 8D resolution. We refer to the 8D product as OceanSODA-ETHZv2.

The paper is organized as follows: First, we outline the data sets and methodology that underlie our innovative technique. Following this, we rigorously evaluate the model's output. We then explore the implications of high-resolution pCO_2 data on FCO_2 variability across different temporal scales. Finally, we assess the local and global impact of these high-resolution estimates, including a case study focused on the influence of a hurricane.

2. Methods

2.1. The OceanSODA-ETHZv2 Method

2.1.1. Design Elements

The method used for the OceanSODA-ETHZv2 product is a classical machine learning-based regression approach to map the sparsely observed fCO_2 data to the global surface ocean. To achieve the mapping at the 8D target resolution, while maintaining robustness and scalability, several design elements are implemented, most of which build on the ideas of other studies.

The first design element concerns the target variable. We first subtract atmospheric CO_2 in order to remove most of the trend in the target variable (Ma et al., 2023). Thereafter, following Bennington et al. (2022), we remove the temperature effect from oceanic fCO_2 leaving the non-thermal, that is, chemically driven, part of the signal. These two transformations capture the impact of two well-understood and quantifiable drivers, such that the machine learning part is focused on the variability imparted on fCO_2 through the other drivers, such as biology and mixing (Sarmiento & Gruber, 2006).

The second design element is that we use an 8-day climatology of ΔfCO_2 as a predictor. This is inspired by the works of Landschützer et al. (2015) and Denvil-Sommer et al. (2019), both of which include information on the fCO_2 seasonal cycle, the dominant mode of variability. The first study does this through clustering based on a predefined monthly climatology of fCO_2, and the second by removing the monthly climatological signal from fCO_2 before training their neural network. We produce this seasonal climatology in a separate initial step.

The third design element is that we combine decision-trees and neural-networks. For the estimation of the climatological seasonal cycle, we use Gradient-Boosted Decision Trees (GBDT), thus taking advantage of the low bias nature of tree-based approaches. For the estimation of the time-varying fCO_2, we use a feed-forward neural network method. Theoretically, the differentiable nature of FFNNs better captures the relationships (i.e., gradients) between the fCO_2 and its drivers (Holder & Gnanadesikan, 2021).

The fourth and final design element is that we include the rate of change of the drivers as predictors. More specifically, the difference between the current and the previous time step is used for variables such as temperature and chlorophyll-a. This gradient adds additional information about the rate of change and further improves the stability of the fCO_2 predictions between time steps.

Overall, this set of design choices improves the stability and robustness of our estimates and improves computational efficiency (compared to other methods), which is important given that the amount of data increased by a factor of ~61 by going from the typical 1M resolution to 8D.

Figure 1 summarizes the 6 steps involved in the estimation of fCO_2 and FCO_2: (a) Data preparation and preprocessing, including the detrending of the target variable by subtracting the atmospheric CO_2 mixing ratio to

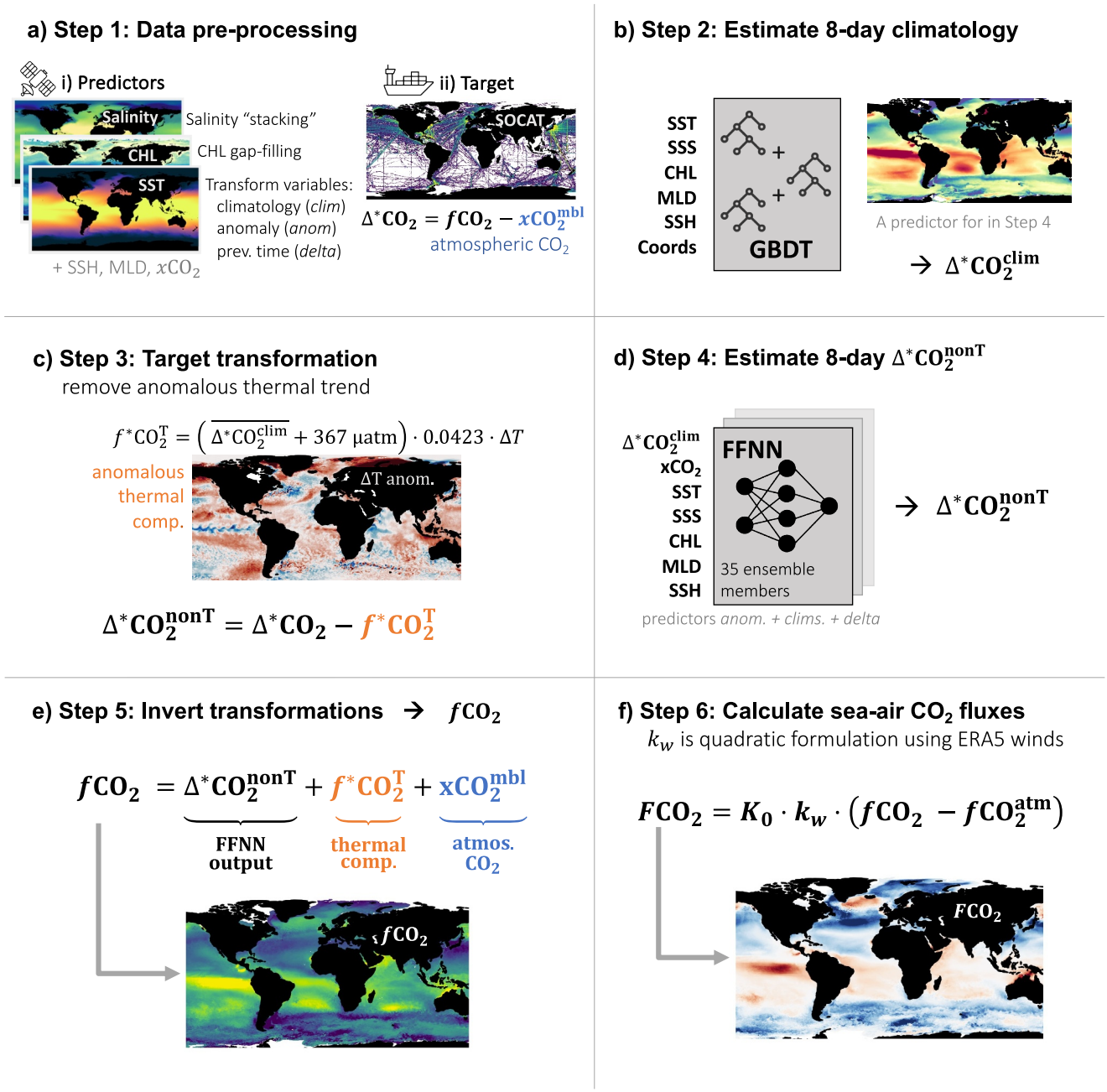

Figure 1. Diagram showing the 6 steps in our approach. (a): Step 1: Data preparation and pre-processing (Figure 2, Section 2.1.2). Notably, we remove the atmospheric CO_2 trend from SOCAT fCO_2 by removing the mole fraction of CO_2 in the marine boundary layer $\left(xCO_2^{mbl}\right)$. This yields the variable Δ^*CO_2. (b) Step 2: An 8-day climatology of Δ^*CO_2 is estimated and used as a predictor in step 4 (Section 2.1.3). To this end, we use Gradient Boosted Decision Trees (GBDT). (c) Step 3: The target variable is transformed by removing the anomalous thermal component, fCO_2^T (after Bennington et al., 2022) (Section 2.1.5). The resulting non-thermal component, $\Delta^*CO_2^{nonT}$, is used as the target for the next step. (d) Step 4: An ensemble of 35 FFNNs predicts $\Delta^*CO_2^{nonT}$ at 8D resolution (Section 2.1.5). (e) Step 5: Reverse transformation: The thermal component (fCO_2^T) and the atmospheric trend $\left(xCO_2^{mbl}\right)$ are added back to arrive at the estimate of fCO_2. (f) Step 6: Calculate the sea-air CO_2 fluxes with the output from the previous step (Section 2.1.7).

form the variable Δ^*CO_2. (b) Machine learning Part 1: Estimation of the 8-day climatology of Δ^*CO_2; (c) Removal of the thermal component from Δ^*CO_2 to form the variable $\Delta^*CO_2^{nonT}$; (d) Machine learning Part 2: Estimation of the time-variable non-thermal target $\Delta^*CO_2^{nonT}$; (e) Reverse transformation of $\Delta^*CO_2^{nonT}$ to obtain the time-variable fCO_2 field; (f) Estimation of the sea-air CO_2 fluxes, FCO_2. The methods section, in large,

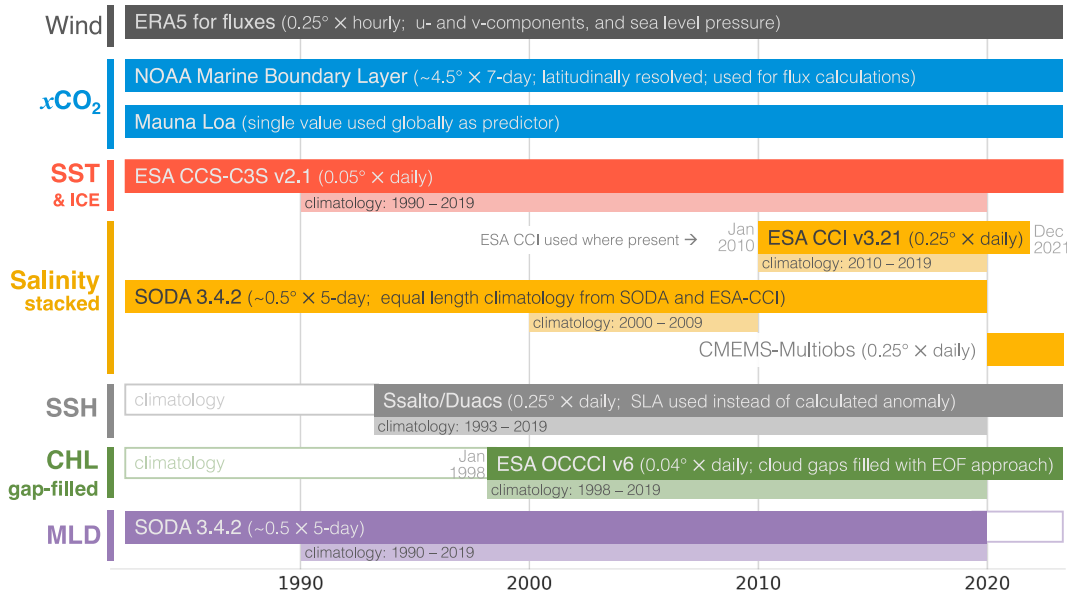

Figure 2. A Gantt chart showing the various data sets used for the OceanSODA-ETHZv2 product. Each variable is represented by a color, with each product having its own horizontal bar and the thinner, lighter shaded bars represent the periods over which the climatology was calculated for that variable. In the case of salinity, where products overlap in time, the high-latitude and coastal gaps of the ESA-CCI product are filled by the CMEMS-Multiobs product, with ESA-CCI always taking priority.

describes these steps in greater detail. Furthermore, we describe the decomposition of the sea-air CO_2 fluxes into different temporal modes of variability.

2.1.2. Step 1: Data Preparation and Preprocessing

We obtain fCO_2 from the SOCAT v2022 cruise track data set (D. C. Bakker et al., 2016; D. C. E. Bakker et al., 2023), retaining data with a quality flag of D or higher. These observations are then binned to a regular grid with a resolution of 8-day by 0.25° × 0.25°. No weighting is applied during the binning, meaning that the data is analogous to the unweighted fCO_2 data from the gridded SOCAT product. We then remove the atmospheric CO_2 component from the data using the 7-day, latitudinally resolved CO_2 concentration in the marine boundary layer $\left(xCO_2^{mbl}\right)$ from Lan et al. (2023):

$$\Delta^* CO_2 = fCO_2^{SOCAT} - xCO_2^{mbl}. \quad (1)$$

For detrending, we use xCO_2 rather than atmospheric fCO_2 to avoid any unwanted impact of variations in atmospheric pressure and vapor pressure (see Equation 4 for the calculation of atmospheric fCO_2).

For the predictors, we use sea surface temperature (SST; Merchant et al., 2019), salinity (Boutin et al., 2018; Carton et al., 2018; Droghei et al., 2016), sea surface height (SSH; Taburet et al., 2019), chlorophyll-a (CHL; Sathyendranath et al., 2023), mixed-layer depth (MLD; Carton et al., 2018), and the atmospheric CO_2 concentration at Mauna Loa (xCO_2^{atm}; Tans & Keeling, 2023) (see Figure 2). An important consideration for the selection of the predictor variables and their sources was their availability at 8D resolution or better. This is achieved for all predictors except for MLD and SSS, which are available only at a resolution of 5 days by 0.5°. Our set of predictors is rather similar to that used by most other gap-filling methods (Chau et al., 2024; Rödenbeck et al., 2015).

Pre-processing and transformation (Figure 1a) of these predictor variables can be separated into: (a) Regridding the predictor variables to the target 8D resolution. For SST, SSH, and CHL this means downgrading the resolution in time from daily to 8-day resolution and in space from ~5 km (SST and CHL). For MLD and SSS, this means downgrading in time (5-day to 8-day), but upgrading in space (from 0.5° to 0.25°). (b) variable stacking for salinity, and gap-filling for CHL, (c) separation of anomalies from climatologies (*clim, anom*), (d) calculating the

difference between the current and previous time step (*delta*) which is used as a predictor. For the full description of how these data were prepared, see Text S1 in Supporting Information S1.

2.1.3. Step 2: Estimation of an 8-Day Climatology

We estimate the quarter-degree 8-day climatology of Δ^*CO_2 using Gradient Boosted Decision Trees. We use seasonal climatologies of SST, SSS, CHL, MLD, and SSH as predictors along with transformations of the time and space coordinates (Figure 2) — see SText S1 in Supporting Information S1. The estimated $\Delta^*CO_2^{clim}$ is smoothed with a rolling mean with a 2-month by 0.75° window (seven 8-day time steps) to avoid overfitting to the subseasonal variability still included in the target. We choose a large temporal window, but narrow spatial window to preserve spatial gradients, particularly in coastal regions.

2.1.4. Step 3: Target Transformation

The target, $\Delta^*CO_2^{nonT}$, is estimated by removing the anomalous thermal component of fCO_2 (Figure 1c), after Bennington et al. (2022):

$$\Delta^*CO_2^{nonT} = \Delta^*CO_2 - \underbrace{\left(\overline{\Delta^*CO_2^{clim}} + 367\ \mu atm\right) \cdot 0.0423 \cdot \Delta T}_{\text{thermal component}}, \quad (2)$$

where, $\overline{\Delta^*CO_2^{clim}}$ is the spatially resolved long-term mean of Δ^*CO_2 from (Step 2) including an offset of 367 µatm (an estimate of the mean oceanic fCO_2 over the study period), and ΔT is the anomaly of sea surface temperature relative to its 8-daily climatology (similar to Bennington et al., 2022). We do not include the seasonal component of temperature, since seasonal information is already included in $\Delta^*CO_2^{clim}$.

2.1.5. Step 4: Estimation of $\Delta^*CO_2^{nonT}$

We use a Feed-Forward Neural Network to estimate $\Delta^*CO_2^{nonT}$. This choice is based on the work of Courtois et al. (2023), who demonstrate that such networks are able to extrapolate beyond training observations. In addition, neural networks were shown to be better able to capture the true relationship between a target variable and its predictors (Holder & Gnanadesikan, 2021).

Predictor variables for this second machine learning step include the climatology of Δ^*CO_2 from step 2, all other predictors (SST, SSS, CHL, MLD, and SSH) as anomalies, together with their 8-day climatologies, and the differences between the current and previous time step. The atmospheric CO_2 trend from Mauna Loa is included as the final predictor.

The 8D binned fCO_2 data and the colocated predictors are split into training, testing, and validation subsets. We use the same approach as Bennington et al. (2022) where every seventh month is considered a test (or validation) month. The validation subset uses the same 7-month split, but with a 3-month offset. The validation split is used to avoid overfitting by the FFNN by stopping training when no improvement is observed. Test data are not used during training of the FFNN and are used to assess performance. Given the structure of our splits, seven train-test-validation splits can be created by starting on a different month in 1982 (January through July), which means that we can fully reconstruct the SOCAT cruise tracks with only test splits (Figure S1 in Supporting Information S1).

We use an ensemble of FFNNs to predict Δ^*CO_2, with five FFNNs per train-test-validation split. Together with the 7 splits, this results in a total of 35 ensemble members. We use TensorFlow and Keras to construct our neural network ensemble (Abadi et al., 2015). Each ensemble member has two hidden layers with 64 and 32 neurons with ReLu activation (see Text S2 in Supporting Information S1).

2.1.6. Step 5: Retransformation

The output of the FFNNs is first transformed back to Δ^*CO_2 from $\Delta^*CO_2^{nonT}$ using the thermal component of Equation 2. The Δ^*CO_2 variable is then further transformed back to fCO_2 using Equation 1. All 35 ensemble members are then averaged for the estimates of fCO_2 and the standard deviation of the ensemble is calculated (σ_{ens}).

2.1.7. Step 6: Calculation of Fluxes

We calculate sea-air CO_2 fluxes (FCO_2) using the bulk formulation:

$$FCO_2 = K_0 \cdot k_w \cdot \left(fCO_2 - fCO_2^{atm}\right) \cdot (1 - ice), \tag{3}$$

where K_0 is the solubility of CO_2 in seawater from Weiss (1974), k_w the gas transfer velocity (R. H. Wanninkhof, 2014; Sarmiento & Gruber, 2006), fCO_2 is the surface ocean CO_2 fugacity predicted by the OceanSODA-ETHZv2 method, fCO_2^{atm} is the atmospheric marine boundary layer CO_2 fugacity, and *ice* is the sea-ice fraction from the temperature product shown in Figure 2. The sign convention is that a positive flux in Equation 3 is upward, that is, indicating outgassing of oceanic CO_2.

Atmospheric fCO_2 is calculated from the dry air mixing ratio of atmospheric CO_2 in the marine boundary layer from NOAA, that is, xCO_2^{mbl} (Lan et al., 2023):

$$fCO_2^{atm} = xCO_2^{mbl} \times (P_{atm} - pH_2O) \times \text{virial factor}, \tag{4}$$

where P_{atm} is the mean sea-level pressure from ERA5 (Hersbach et al., 2020), pH_2O is the partial pressure of water vapor based on Weiss and Price (1980), and the *virial factor* accounts for the non-ideal behavior of CO_2 (Weiss, 1974). For k_w, we use the quadratic formulation of the sea-air CO_2 fluxes from R. H. Wanninkhof (2014) scaled for ERA5 winds:

$$k_w = 0.274 \cdot U_{10}^2 \cdot \left(\frac{Sc}{660}\right)^{-1}, \tag{5}$$

where U_{10}^2 is the second moment of the wind speed, and Sc is the Schmidt number for CO_2 (Jähne et al., 1987). The second moment of the wind speed is calculated from the hourly ERA5 wind speed, which is not available in the output but can be estimated from the wind vector components u and v with $u^2 + v^2$. U_{10}^2 is then regridded in the time dimension to match the resolution of our output. The coefficient of gas transfer (0.274) was obtained by ensuring that the global mean gas transfer coefficient for the period 1990 and 2019 and for the ice-free ocean matches the constraint of $k_w = 16.5 \pm 3.2$ cm hr^{-1} (Fay et al., 2021; Naegler, 2009; Sweeney et al., 2007).

2.2. Decomposition of Fluxes

We decompose the temporal variability of the sea-air CO_2 fluxes into three modes by integrating the frequencies in the Fourier domain (after Gu et al., 2023): subseasonal, <3 months frequencies; seasonal, 3 to ~15 months frequencies; and interannual, >15 months frequencies. In some cases, we further separate the interannual variability into sub-decadal (15 months–8 years) and decadal (>8 years) variability. For the Fourier decomposition, we assume that all time steps have the same length, that is, $\frac{365}{46}$ days for the high-resolution 8D product and $\frac{365}{12}$ days for the 1M low-resolution product.

To identify the drivers of the sea-air CO_2 flux variability, we apply a Reynolds decomposition $\left(y = \bar{y} + y'\right)$ to Equation 3. We thereby combine k_w and K_0 as the gas transfer coefficient $k_x = K_0 \cdot k_w$, in order to focus on the role of the wind variability (Doney et al., 2009)—the temperature dependencies in k_w and K_0 account for <1% of the variability of k_x (Woolf et al., 2016). We also look only at the role of the sea-air difference in fCO_2, that is, ΔfCO_2, since the variability of atmospheric fCO_2 is much smaller than that of the oceanic fCO_2. Furthermore, we neglect the role of sea-ice variations. With these simplifications, we decompose the sea-air CO_2 flux, FCO_2, as follows:

$$FCO_2 = \underbrace{\overline{k_x} \cdot \overline{\Delta fCO_2}}_{\text{mean state}} + \underbrace{k_x' \overline{\Delta fCO_2}}_{\text{wind variability}} + \underbrace{\overline{k_x} \Delta fCO_2'}_{fCO_2 \text{ variability}} + \underbrace{k_x' \cdot \Delta fCO_2'}_{\text{cross-term}}. \tag{6}$$

Since we are interested in the variability of FCO_2 we do not have to consider the first term where the long-term averages are taken for both variables. We calculate the variability as representations of the variance (σ^2). If we represent each of the terms in Equation 6 as a, b, and c respectively, the total variance is given by:

$$\sigma^2(a+b+c) = \sigma^2(a)+\sigma^2(b)+\sigma^2(c) + \underbrace{2\cdot\mathrm{cov}(a,b)+2\cdot\mathrm{cov}(a,c)+2\cdot\mathrm{cov}(b,c)}_{\text{covariances} \approx \text{residual}}. \quad (7)$$

Given that we calculate the variance from the frequency domain, we calculate the covariance for the different modes of variability as the residual of the total and summed variances. The fraction contribution by each term is computed as:

$$\frac{\sigma^2(\text{wind})}{\sigma^2(\text{total})}, \frac{\sigma^2(f\mathrm{CO_2})}{\sigma^2(\text{total})}, \frac{\sigma^2(\text{cross term})}{\sigma^2(\text{total})}, \frac{\text{covariances}}{\sigma^2(\text{total})}. \quad (8)$$

Note that since the covariances can be negative, the fractional contribution can also be negative.

3. Evaluation and Assessment

We first evaluate the 8D $f\mathrm{CO_2}$ product estimated by OceanSODA-ETHZv2 by determining the large-scale offsets against the test and training data, and by comparing these offsets against the ensemble spread that we use as an estimate of the prediction uncertainty. We then investigate in what way the high-resolution product is able to capture finer-scale structures in time and space. To this end, we focus on high-frequency observations from open ocean and coastal mooring stations (Sutton et al., 2019). We also assess how well the high-resolution estimates can track high-resolution spatial features observed along cruise tracks.

3.1. Uncertainties

To determine the bias and root mean squared differences (RMSD) of the 8D product against the SOCAT data, we rely only on predictions that have not been used to train the subset of ensemble members. Given that we use an ensemble of results where the starting month changes (1–7), we have a complete representation of the SOCAT data set (Figure S1 in Supporting Information S1; Gregor et al., 2019; Bennington et al., 2022).

The unweighted bias and RMSD for the $f\mathrm{CO_2}$ are low at a global scale at −0.06 µatm and 19.2 µatm, respectively (Table S2 in Supporting Information S1), where a negative bias indicates that OceanSODA-ETHZv2 underestimates $f\mathrm{CO_2}$ relative to SOCAT. The open (coastal) ocean has a bias of 0.07 µatm (−0.23 µatm) and an RMSD of 13.0 µatm (25.4 µatm) (Figure 3, Table S2 in Supporting Information S1), where the coastal ocean is defined as the ocean region within 300 km from the coast or the 1,000 m isobath (Laruelle et al., 2017; Resplandy et al., 2024). Given the spatial and seasonal inhomogeneity in the SOCAT observations, biases and RMSDs are also assessed over time and space (Figure 1 with weighted averages). We assess both metrics for three latitude bands: the high northern (blue lines; >35°N) and southern latitudes (yellow lines; >35°S) and the bounded lower latitudes (red).

Biases are larger and more variable before 2000 when the observations are more sparse (Figure 3a). The spatial separation also reveals that biases compensate meridionally, resulting in lower global biases due to aggregation. There is no seasonality in the biases in any of the latitude bands. Spatially (Figure 3e), biases are mostly low ($|bias| < 2.5$ µatm), except for the coastal regions and a few open ocean regions, for example, the southern Indian Ocean, where there is a positive bias (red in Figure 3e).

In northern high latitudes, RMSD increases slightly over time (∼15 to >20 µatm; blue line in Figure 3c). This is likely due to the increase in the variance of the training data over time, particularly in the coastal ocean, where there is typically higher biogeochemical complexity (Figure S2). The RMSD for the tropics and southern regions remains constant (∼15 µatm) throughout the 40-year period, with no significant slope. On average, RMSDs are larger in the summer months compared to the winter months (∼25 µatm vs. ∼15 µatm) of both the northern and southern high latitudes (blue and yellow lines in Figure 3d). The seasonality of the RMSD in the low latitudes (red line in Figure 3d) is lower by comparison, but has a slight bias to the Northern Hemisphere summer, likely due to sampling bias.

At a global scale, the spatial distribution of the RMSD resembles the ensemble standard deviation (σ_{ens}) with a spatial correlation r^2 of 0.65 for the time averaged maps (Figure 4). This suggests that the ensemble spread is a

10.1029/2024GB008127

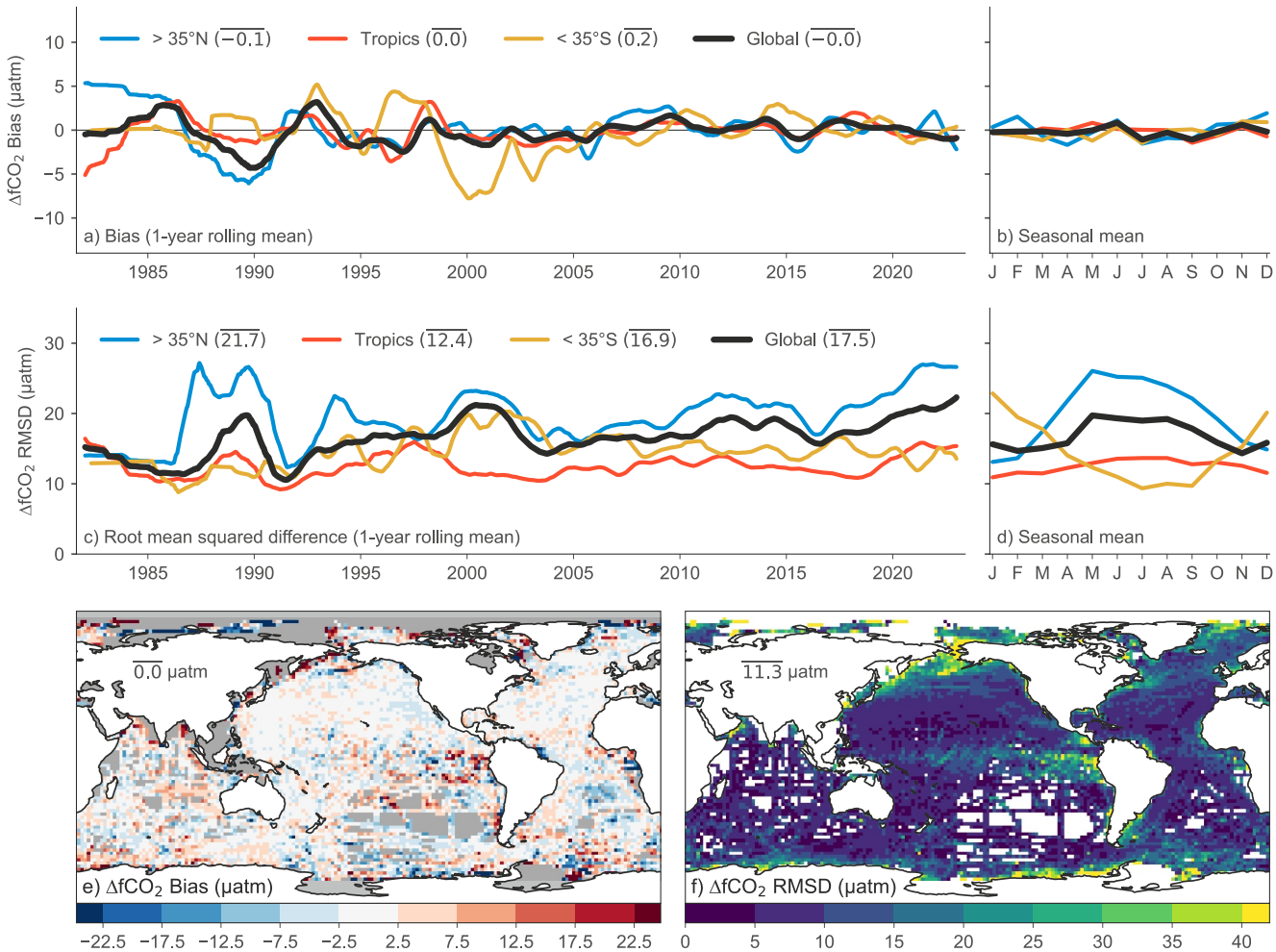

Figure 3. Metrics using test-data estimates (unseen by the machine learning model). Time series on the left show latitudinally averaged biases (a) and root mean squared difference (RMSD in c) with a 1-year rolling mean applied. The legend also contains the time-averaged mean values (in µatm) weighted by number of samples. The smaller time series figures on the right show seasonally averaged biases (b) and RMSD (d). The maps below show test bias (e) and RMSD (f) with the values showing the spatial average of the respective metrics weighted by area.

good indicator of the spatial structure of the quality of the estimated fCO_2. This relationship was exploited by Chau et al. (2024) who scaled σ_{ens} to match the RMSD as closely as possible. Here, we follow the same idea to obtain a global map of uncertainty of our 8D product, but we simply multiply σ_{ens} with a factor of 5.7 (Figure 4), that is, the global mean ratio of $\frac{\overline{RMSD}}{\overline{\sigma_{ens}}} = \frac{11.3}{2.0}$.

3.2. Assessing Spatial and Temporal Scales of Variability

The benefit of estimating high-resolution fCO_2 becomes most obvious when comparing the estimated product against the raw observations in SOCAT. The majority of the SOCAT data set consists of ship cruise tracks. In Figure 5, we compare three representative cruise tracks in the Atlantic and Pacific with the OceanSODA-ETHZv2 test-subset of fCO_2, where the test-subset is a subset of our output that has not been trained with the data it estimates. Along all three tracks, the large-scale variability of the SOCAT fCO_2 observations is well represented. In open ocean regions, the RMSD scores are actually substantially lower than the global average. An exception to this occurs in poorly sampled regions, such as the South Atlantic gyre (Figure 5d, where the regional biases can be >|5| µatm.

Most important is that our high-resolution product captures a substantial fraction of the fine-scale structures contained within the observed fCO_2. For example, the 8D product properly represents the strong outgassing signals associated with the upwelling driven by the gap winds off the coast of Central America (Figure 5a). The

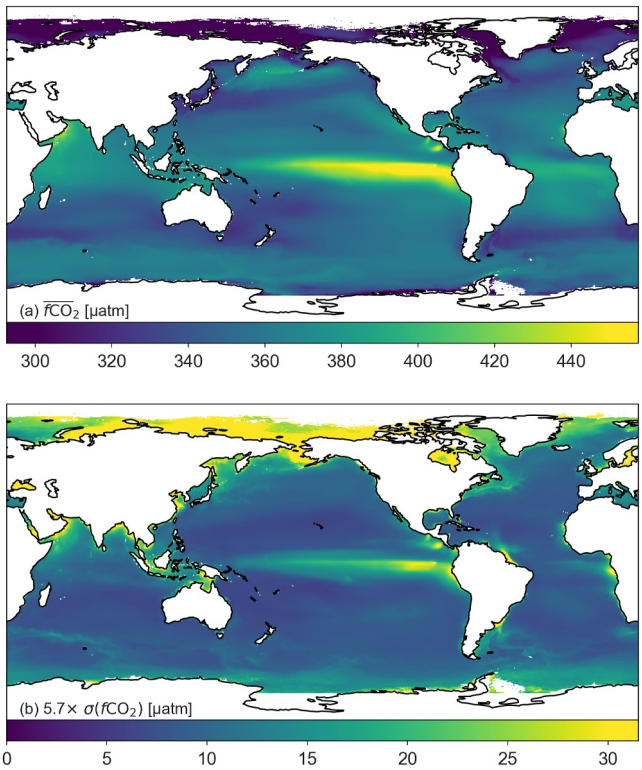

Figure 4. (a) The long-term mean of fCO_2 for 1982 to 2022 and (b) the scaled standard deviation (5.7σ) of the 35 ensemble members.

impact of these winds, the Tehuano and Papagayo jets, on the surface ocean CO_2 system on ocean biogeochemistry can be clearly seen, extending in the case of the Papagayo jet all the way into the region of the Costa Rica dome (Chapa-Balcorta et al., 2015; Fiedler, 2002). These features are generally not well captured by the global 1M products (Figure S4 in Supporting Information S1). Further, the 8D product adequately captures the magnitude of the equatorial upwelling in both the Atlantic and Pacific (Figures 5c and 5d), as well as the strong variations off the coast of Europe (Figure 5d).

At the same time, there are clear limits in our 8D product. The specific structure of the upwelling signal associated with the gap winds is missed, as is the exact location of the equatorial upwellings offset. Furthermore, the signals of two tropical instability waves in the equatorial Pacific (Figure 5c) are completely missed by the 8D product. This is largely a consequence of these features having a propagation speed of around 30 km day^{-1} (Legeckis, 1977), such that an 8-day resolution ~25 km product is insufficient to correctly capture their dynamics. Finally, the neural network cannot capture some extremely sharp gradients observed near New Zealand (Figure 5c). However, our 8D product captures the observed variations in fCO_2 much better than the 1M products (see Figure S4 in Supporting Information S1).

Similar successes and limitations of the 8D product can be identified when comparing it with high-frequency observations from long-term moorings. In Figure 6, we show comparisons of our 8D product with observations from four open-ocean and one coastal mooring programs in the Pacific and the Pacific sector of the Southern Ocean (Sutton et al., 2019; D. C. Bakker et al., 2016). The observed fCO_2 is well represented by the estimated fCO_2 at locations with a large seasonal cycle. The 8D product captures 66% of the variability observed at the coastal California Current Ecosystem 1 (CCE1) mooring 140 mi (ca. 225 km) off the coast of California (Figure 6a), 82% at the SOFS mooring south of Tasmania (Figure 6e) and 72% at the STRATUS mooring site off the coast of South America (Figure 6f). Also, shorter-term variations are generally well captured at these three sites. The performance of the 8D product is somewhat weaker at the other two sites. While the fCO_2 estimates at the PAPA mooring location (Figure 6b) are able to represent most of the seasonal variability, some extremes are not captured. This leads to only 50% of the observed variability being captured by the 8D product. In the equatorial Pacific (TAO125W in Figure 6d) the 8D product appears to miss a substantial fraction of the high-frequency variability. Much of this variability is caused by equatorial instability waves. This mismatch is consistent with the mismatches seen in the Equatorial Pacific when analyzing the cruise-line data (Figure 5c).

Comparisons with a larger set of moorings reveal similar strengths and limits (Table S4). For example, our 8D product reproduces the observed variations at the WHOTS mooring near Hawaii exceptionally well. But it struggles to capture the variations at the very coastal CCE2 mooring off the coast of California. This follows the previously identified pattern, with our 8D product reproducing the observations best in regions where the seasonal cycle dominates, while it performs less successfully in regions where high-frequency variations dominate. This confirms our previous findings that while our product made a substantial step forward in reproducing higher frequency variations, it is still too coarse to capture all of them. This is also in part due to the fact that our method has a propensity to adhere to the seasonal cycle, which is very helpful overall, as this allows us to capture the dominant signal of variability over most of the ocean. Although this increases the robustness of the method, it comes at a cost.

4. Patterns and Variability of fCO_2

4.1. Representation of High-Resolution Features

Recognizing some shortcomings of our high-resolution mapped fCO_2 product, it is instructive to visualize its strengths in representing fine-scale features previously not seen in gap filled fCO_2 products. In Figure 7 we show five snapshots from different regions of the global ocean (Movie S1). The 8D estimates are able to represent a lot

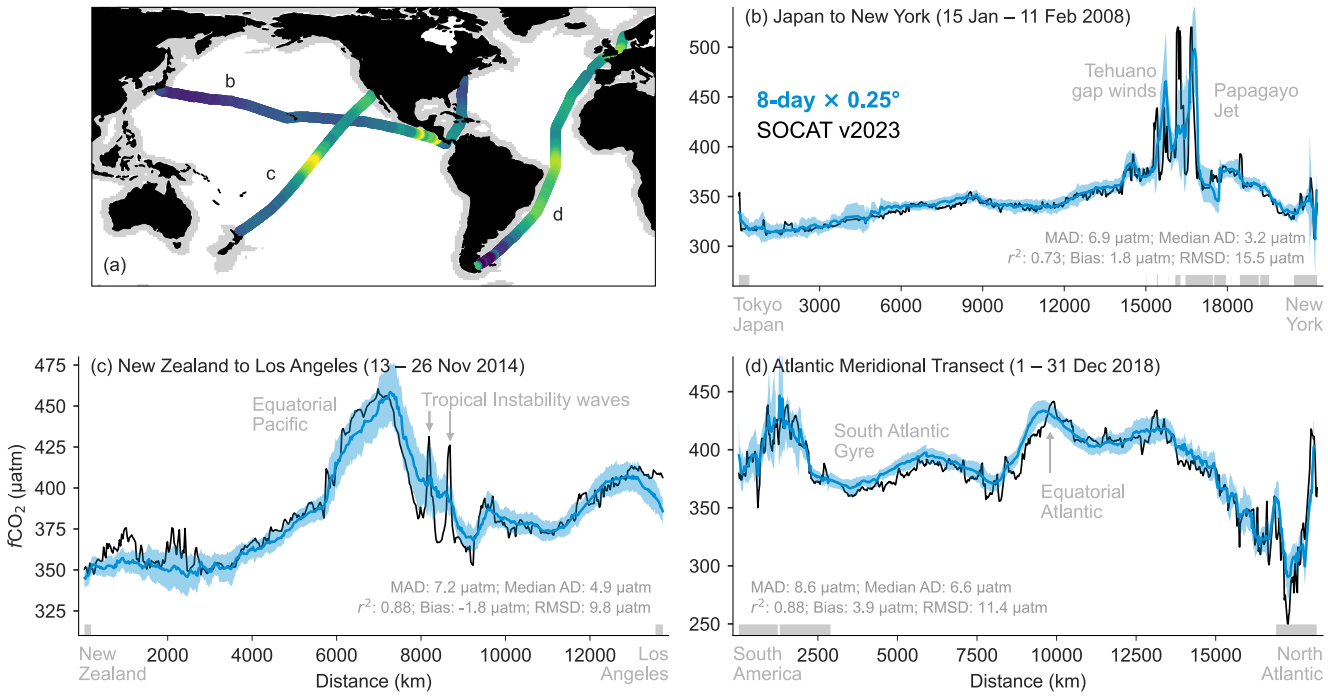

Figure 5. Comparison of the observed fCO_2 along three selected cruise tracks contained within the SOCAT database (resampled to 8D; black lines) with the estimated fCO_2 from the 8D test-subset (i.e., unseen by the model; blue lines). The blue envelope shows the scaled ensemble standard deviation ($5.7\sigma_{ens}$). (a) Locations of the cruise tracks with the colors indicating the measured fCO_2. The gray region indicates the coastal mask used in this study from Resplandy et al. (2024). (b) Transect from Japan to New York occupied between 15 Jan 2008 and 11 Feb 2008 (SOCAT cruise: 49P120080116, investigator: Nojiri, Y). (c) Transect from New Zealand to Los Angeles occupied between 13 Nov 2014 and 26 Nov 2014 (SOCAT cruise: AG5W20141113, investigators: Cosca, C., Feely, R., Alin, S.). (d) Transect between Southampton and Punta Arenas occupied between 1 Dec 2018 and 31 Dec 2018 as part of the Atlantic Meridional Transect program (SOCAT cruise: 58US20181201, investigator: Fransson, A). The distance along the *x*-axes of the cruise tracks (b, c, d) is plotted with zero being the westernmost point, regardless of the actual direction of travel. Gray bars along the *x*-axis indicate where cruise tracks are in the coastal area shown in (a).

of fine-scale spatial variability in fCO_2 that could not be seen so far in the 1M products. For example, the 8D estimates depict important fCO_2 details in the eastern boundary upwelling regions off northwestern Africa (Figure 7a) and off the U.S. West Coast (Figure 7b). Of special note are the high fCO_2 values tagging the coasts, reflecting recently upwelled waters, and the rapid offshore decrease of fCO_2 primarily due to strong biological drawdown. In addition, the filamentous features of these low fCO_2 waters are clearly visible. These structures correspond very well to detailed regional observations and modeling studies (Friederich et al., 2002; Lachkar & Gruber, 2013; Turi et al., 2014).

Similarly, off the West coast of Central America, the 8D estimates reveal the spatial extent of the upwelling-driven maxima in fCO_2 downstream of the Papagayo and Tehuantepec gaps (Figure 7c). Despite the ephemeral nature of these gap winds, which last for hours to several days (Liang et al., 2009; Romero-Centeno et al., 2003), their oceanic signatures persist long enough to be picked up well by our 8 daily product. The spatial structure of these mountain gap wind features in fCO_2 is consistent with what is known from in-situ observations (see comparison to the cruise data above (Figure 5a) and the work by for example, Chapa-Balcorta et al. (2015). Still, the spatial mismatches we identified in comparison with the cruise-line data suggest that the 8-day resolution is not entirely sufficient to fully capture these ephemeral events.

The high spatial resolution also can resolve the very dynamic structure of fCO_2 in the Malvinas Current region (Figure 7d). In this region, strong biological productivity on the Patagonian shelf interacts in a complex manner with the mixing of very different waters masses, that is, the mixing of the warm southward-flowing Brazil Current (BC) and the cold northward flowing Malvinas Current (MC) (Arruda et al., 2015). The meandering features have been seen in modeling studies (Arruda et al., 2015) and the strong gradients created by the mixing are also regularly captured in the raw SOCAT cruise data (see also Figure 5d).

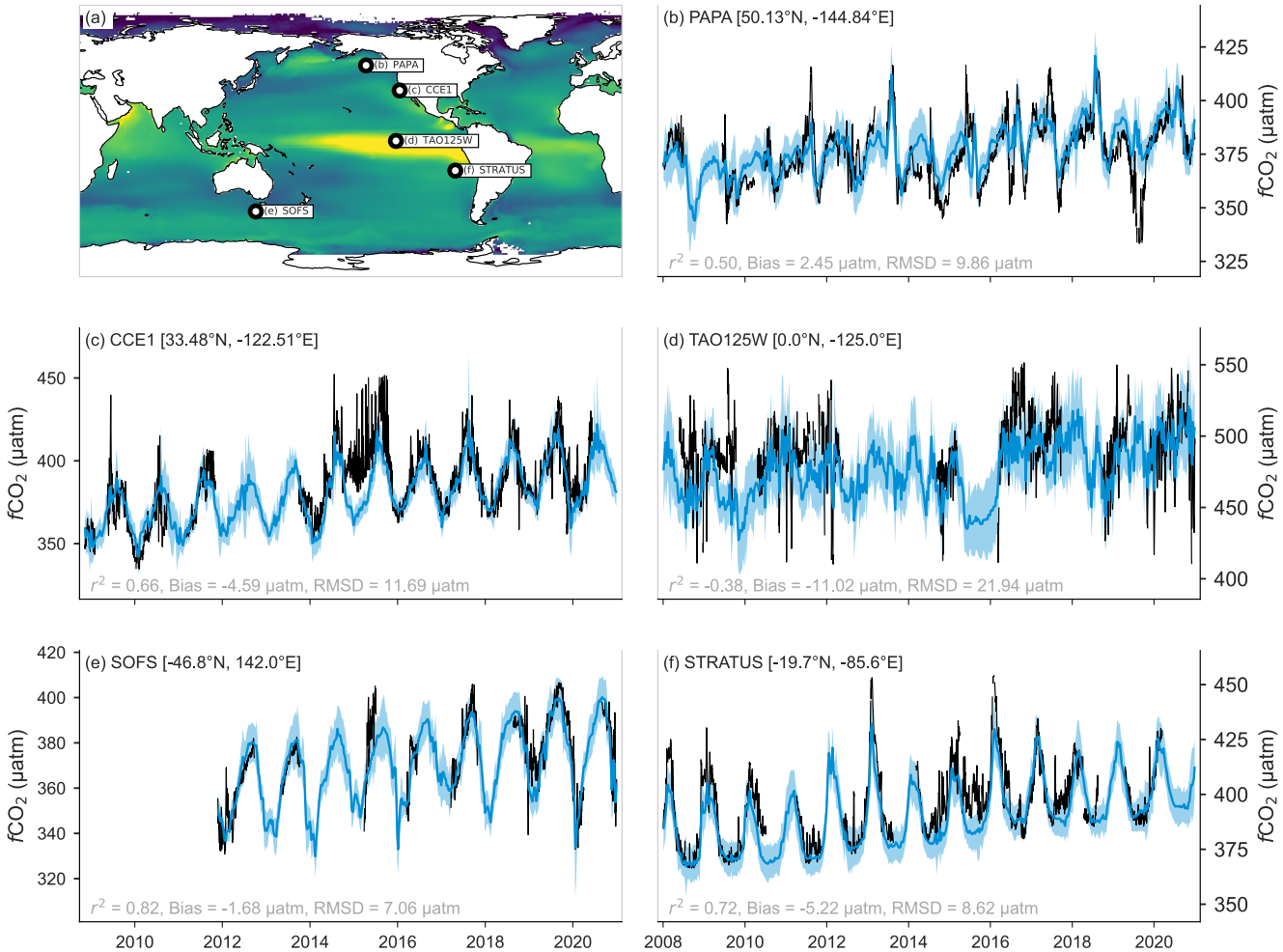

Figure 6. Comparison of the OceanSODA-ETHZv2 estimated fCO_2 with observations from 4 open ocean and 1 coastal mooring stations. (a) Map of the long-term mean fCO_2 with the location of the mooring sites. (b) Comparison at the PAPA mooring site in the Subarctic Pacific. (c) Comparison at the CCE1 coastal site in the California Current System. (d) Comparison at the TOA site within the central tropical Pacific. (e) Comparison at the SOFS site south of Tasmania. (f) Comparison at the STRATUS mooring site off the coast of Chile. Blue lines show the "unseen" test-subset of estimated ΔfCO_2, with the blue envelope showing the scaled ensemble standard deviation ($5.7\sigma_{ens}$). Black lines show the measured mooring data resampled to a daily resolution. Model metrics are shown in gray in the bottom of each plot. Note that these estimates are not used to train the model.

And finally, in the equatorial Pacific, the 8D fCO_2 product begins to resolve the tropical instability waves (Figure 7e). However, their shapes are slightly distorted and overly smoothed, which is expected given their rapid propagation speeds (Legeckis, 1977). This mismatch has already been seen in the cruise-line data (see also Figure 5c) and also the mooring data from the TAO125W site (Figure 6d).

4.2. Comparison With Other fCO_2 Products

The improvement of the 8D fCO_2 estimates is also evident when comparing it to several other fCO_2-products (Figure 8), namely, CMEMS-LSCE (1M by 0.25° Chau et al., 2024), ULB-MPI-SOMFFN (1M by 0.25° Roobaert et al., 2023), Jena-CarboScope by (daily by 2.0° Rödenbeck et al., 2014), MPI-SOMFFN by (1M by 1° Landschützer et al., 2016), and LDEO-HPD by (1M by 1° Bennington et al., 2022). Apparent in the comparison is that the fine-scale gradients of the 8D OceanSODA-ETHZv2 estimates are sharper compared to the other approaches. In this scenario, we find that the mean gradients $\left(\overline{\|\nabla fCO_2\|}\right)$ of OceanSODA-ETHZv2 fCO_2 are 60% stronger than the other two high-resolution products (Figures 8b and 8c), and more than three times those of the monthly by 1° × 1° resolution products (Figures 8e and 8f). The gradient increases to 0.41 µatm km^{-1} when considering the 95th percentile (representing dynamic regions) or ∼8 µatm over 20 km, thus approaching the sharp gradients recorded in observational studies (i.e., ∼10 µatm over 20 km Sutton et al., 2021).

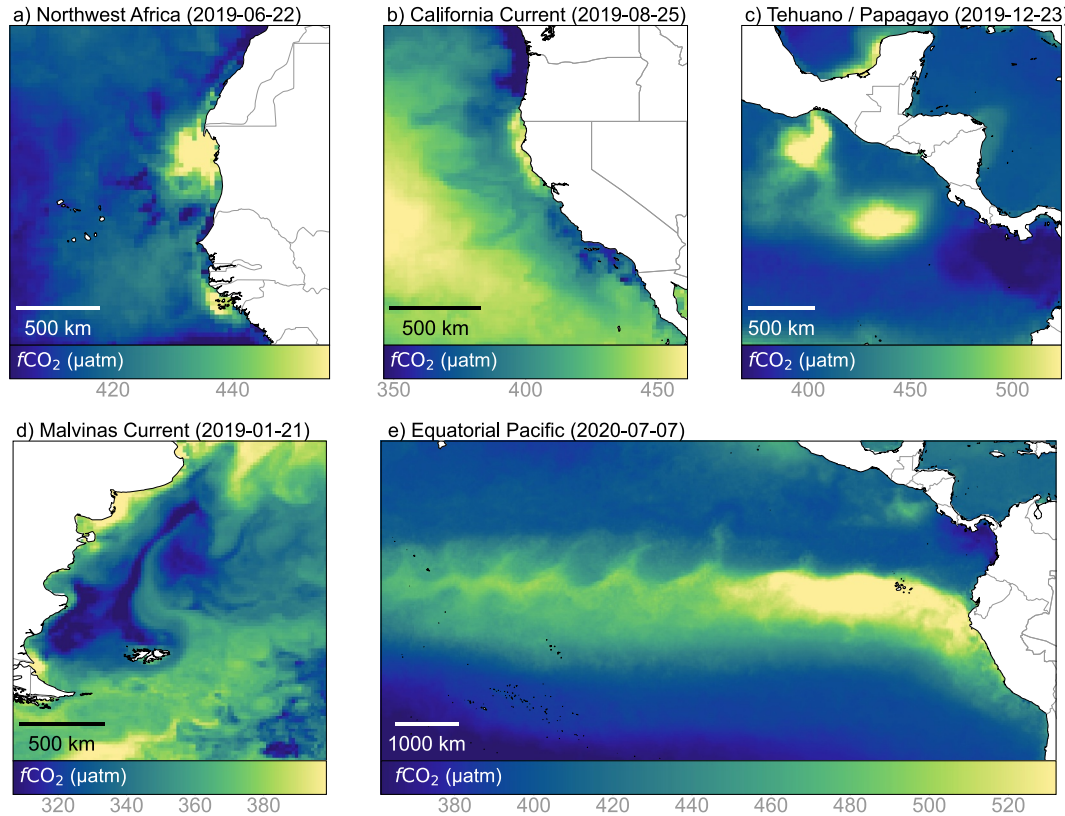

Figure 7. fCO_2 for different regions with the period represented shown by the date in brackets. (a) Northwest African coastline, (b) The Northwest coastline of the USA, with eastern boundary upwelling occurring in (a) and (b). (c) The West coast of Central America where the Tehuano gap winds, and Papagayo Jet winds occur, leaving the imprint of the Costa Rica Dome, (d) the Malvinas (Falkland) current off the southeast coast of Argentina, (e) the equatorial Pacific from 165°W to 75°W.

In summary, our method succeeds in capturing fine-scale spatial variability of fCO_2 in regions where gradients in driver variables are strong. At the same time, there are still high-frequency features (e.g., tropical instability waves, coastal dynamics) that are not well captured, even at the 8D resolution. However, our method is capable of capturing realistic variability that other methods tend to underestimate.

5. Variability of the Ocean Carbon Sink

5.1. Mean, Trend, and Variability of FCO_2

The globally integrated sea-air CO_2 flux FCO_2 inferred from the OceanSODA-ETHZv2 product decreases (but increases in strength) from -1.5 PgC yr^{-1} in 1990 to -2.3 PgC yr^{-1} in 2019, with a mean uptake of -1.92 PgC yr^{-1} over the 1990 through 2019 period. This FCO_2 estimate falls within the 1σ-bounds of the SeaFlux ensemble of six fCO_2-products of -1.92 ± 0.20 PgC yr^{-1} calculated with ERA5 winds (Table 4 in Fay et al., 2021). Accounting for the outgassing of natural carbon associated with the balance between river input and burial, that is, the so-called steady-state river outgassing flux of about 0.65 PgC yr^{-1} (Regnier et al., 2022), the OceanSODA-ETHZv2 product implies a total sink for anthropogenic CO_2 of -2.33 PgC yr^{-1} for 1990–1999, of -2.52 PgC yr^{-1} for 2000–2009, and of -2.85 PgC yr^{-1} for 2010–2019. This is at the lower end of the uptake estimates summarized by (Gruber et al., 2023) for the three decades using observations, but is higher than most ocean biogeochemical models suggest (Hauck et al., 2020). Concerning the FCO_2 trend, OceanSODA-ETHZv2 estimates -0.26 PgC yr^{-1} decade^{-1} over the period 1990–2019 (Figure 9a). For the period used in RECCAP2 (2001–2018; DeVries et al., 2023), we estimate and FCO_2 trend of -0.33 PgC yr^{-1} decade^{-1}, which is lower than the RECCAP2 ensemble of fCO_2-products (-0.61 ± 0.12 PgC yr^{-1} decade^{-1}), but nearly the same as the trend for the RECCAP2 biogeochemical model ensemble (-0.34 ± 0.06 PgC yr^{-1} decade^{-1}). Concretely, our approach, while

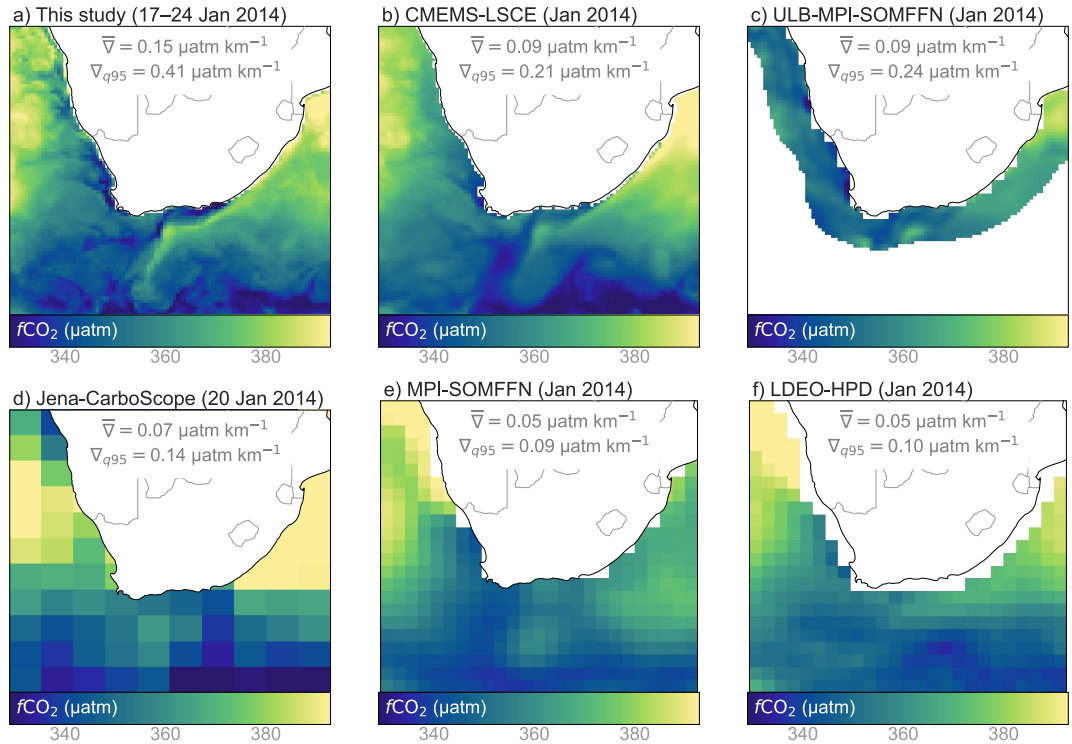

Figure 8. A comparison of fCO_2 from different mapping products around Southern Africa, which includes the Agulhas Current and Benguela upwelling system. (a) OceanSODA-ETHZv2 (this study) for the 8-day period 17–24 January 2014 (8-day by 0.25°). (b) CMEMS-LSCE (monthly by 0.25°) by Chau et al. (2024), (c) ULB-MPI-SOMFFN (monthly by 0.25°, coastal only) by Roobaert et al. (2023), (d) Jena-CarboScope (daily by 2°) by Rödenbeck et al. (2022), (e) MPI-SOMFFN (monthly by 1°) by Landschützer et al. (2016), and (e) LDEO-HPD (monthly by 1°) by Bennington et al. (2022). The metric $\overline{\nabla}$ represents the average horizontal gradient of fCO_2 for the plotted region and ∇_{q95} represents its 95th percentile.

on the conservative side, is consistent with the current best estimates of the magnitude and change of the ocean carbon sink (DeVries et al., 2023; Gruber et al., 2023).

The globally integrated flux varies substantially around these mean uptakes, with the seasonal cycle contributing the most variance (Figure 9b). There is also evidence of an increase in the magnitude of the seasonal cycle over time, confirming previous findings based on theory, models, and observations (Landschützer et al., 2018; Rodgers et al., 2023). In addition, clear subseasonal variations are visible in the global time series, as well as interannual to decadal variations. Compared to other gap filled fCO_2 products (DeVries et al., 2023), the interannual to decadal variability of the OceanSODA-ETHZv2-based flux estimate tends to be on the lower end of the spectrum. Still, it shows a clear stalling of the trend toward stronger uptake in the 1990s, and an acceleration thereafter (DeVries et al., 2019; Gruber, Clement, et al., 2019; Gruber et al., 2023; Landschützer et al., 2015).

To assess the impact of higher resolution on our estimates of the variability of the ocean carbon sink, we contrast our 8D estimate with a monthly by 1° × 1° (1M) estimate we obtain by averaging fCO_2 from OceanSODA-ETHZv2 and all other inputs to Equation 3 to 1M (Figure 9a). Although the 1M results are based on the averages of the 8D estimates, the global fluxes of the two products are not identical on longer timescales, that is, the 1M product has a marginally smaller mean (1990–2019) uptake of −1.87 PgC yr^{-1}. This small difference is likely due to small covariances between wind and ΔfCO_2, whose magnitude is scale-dependent (see below). As expected, much larger differences occur on subseasonal timescales, where the 8D product reveals higher crests and deeper troughs than the 1M product. This results in higher temporal variance, which is best analyzed spatially.

The map of the total standard deviation of the sea-air CO_2 flux FCO_2 reveals strong spatial differences ranging from near zero to more than 7 mmol m^{-2} day^{-1} (Figure 9b). The northern mid-to high-latitudes have the highest variability (>5 mmol m^{-2} day^{-1}), particularly in the Pacific basin. Some island and coastal regions (e.g., the Kerguélen Plateau and Oman upwelling regions) have similarly high variability. The low-latitude regions have

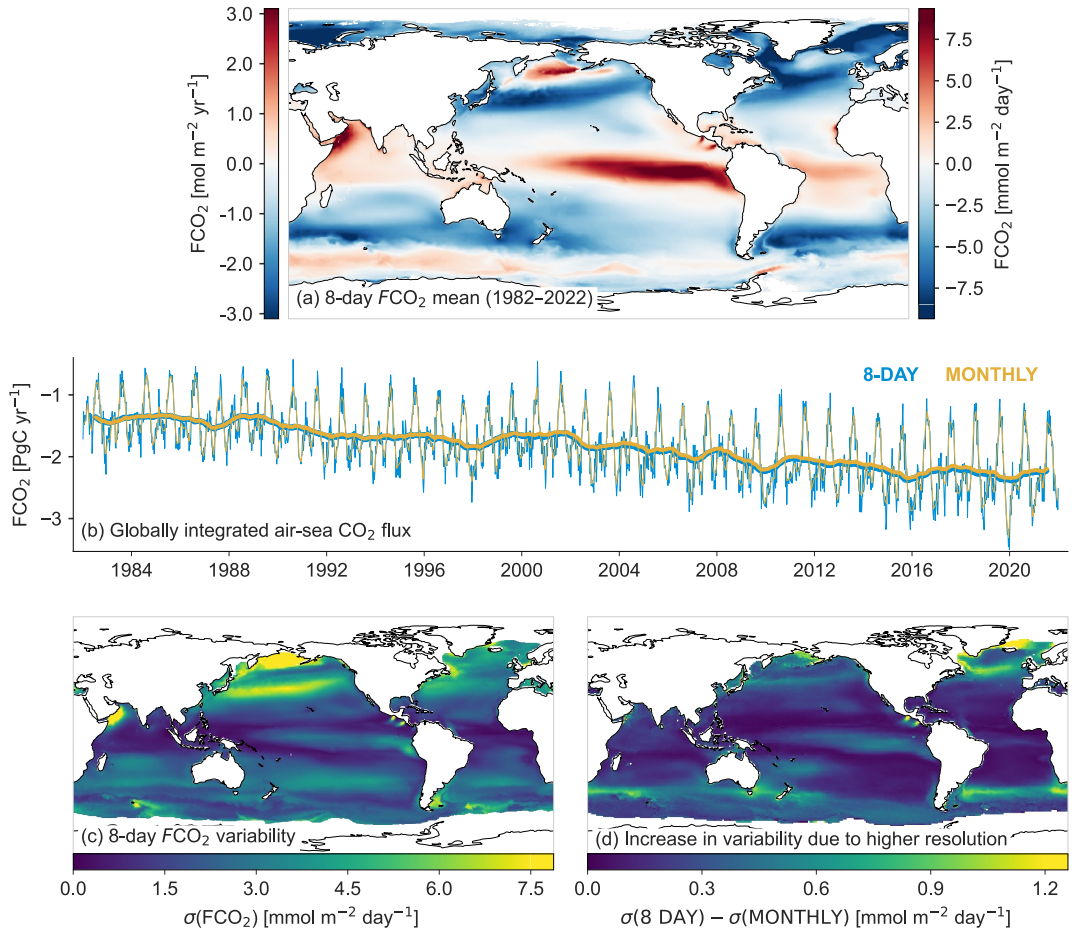

Figure 9. Analysis of the OceanSODA-ETHZv2 sea-air CO_2 fluxes (FCO_2) (negative means ocean uptake). (a) Map of the global long-term mean sea-air CO_2 flux density (1982–2022). The color bars on either side of the figure show different units as indicated. (b) Time series of the globally integrated sea-air CO_2 fluxes (FCO_2) for 8-day (8D, blue) and monthly (1M, yellow) estimates, where the latter is calculated by resampling the inputs of Equation 3 to a lower monthly × 1.0° resolution and then calculating FCO_2. (c) The standard deviation (σ) of FCO_2 for the 8D estimates. (d) The difference between (c) and the resampled 1M (monthly × 1.0° resolution).

low variability (<1 mmol m^{-2} day^{-1}), except for the central and eastern tropical Pacific. In the global mean, the temporal standard deviation of the 8D product amounts to $\overline{\sigma}$ = 2.45 mmol m^{-2} day^{-1}. This is about 10% more than the global mean temporal standard deviation $\overline{\sigma}$ of the 1M product (2.20 mmol m^{-2} day^{-1}).

Although the global increase in the captured variability of the 8D product is modest, regionally, the increase in the variability of the sea-air CO_2 fluxes can be much more pronounced (Figure 9c). The increase in variability is most notable along a band (~40°S) in the Sub-Antarctic Zone (SAZ) in the Southern Ocean, where the 8D product increases the variability relative to the 1M product by more than 50% in some parts. This corresponds well with the results from (Monteiro et al., 2015). The northern high-latitude Atlantic Ocean exhibits a particularly large increase in variability (≳1.5 mmol m^{-2} day^{-1}). However, the relative increase is smaller compared to the SAZ (<40%).

5.2. Temporal Decomposition of Variability of FCO_2

As seen already in the global time series, the total variability of FCO_2 is dominated by the seasonal mode (10b), that is, by variability on timescales between 3 and 15 months. The standard deviation of the flux on this timescale goes up to more than 6 mmol m^{-2} day^{-1}, with the highest values found in the North Pacific. In contrast, in the tropics, variability on seasonal timescales is very low. Globally, the standard deviation on seasonal timescales

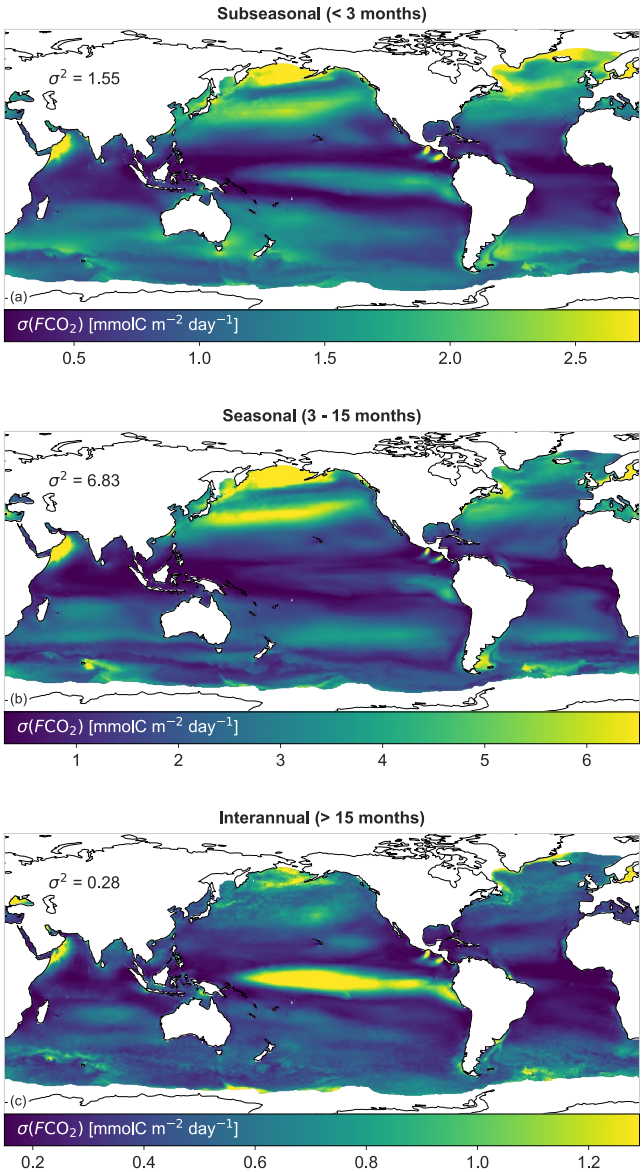

Figure 10. Maps of the standard deviation of the seasonal (a), subseasonal (b), and interannual (c) components of FCO_2 from the 8D product of OceanSODA-ETHZv2. Values on the plots show the area weighted mean of the variance in (mmol m^{-2} day^{-1})^2. See Tables S3, S4, and S5 in Supporting Information S1 for low and high-resolution variance, standard deviation, and percentages, respectively.

amounts to σ = 2.61 mmol m^{-2} day^{-1}. The difference between σ(1M) and σ (8D) is small (0.03 mmol m^{-2} day^{-1}), indicating that, on seasonal scales, monthly data capture most of the variability.

The subseasonal mode (<3 months) is the next largest contributor to FCO_2 variability (σ = 1.24 mmol m^{-2} day^{-1}, Figure 10a). The contribution to variability on these shorter timescales is dominated by the mid-latitudes and the equatorial Pacific, with some regions exceeding 2.5 mmol m^{-2} day^{-1}. To first order, the spatial distribution of the subseasonal variability is similar to that of the seasonal variability (Pearson R = ~0.69). An important difference is the strong subseasonal contribution in the Oman upwelling system and in the Coral Sea and Tasman Sea regions east of Australia. As expected, the gain in information by going from 1M to 8D is the largest in this subseasonal mode. Globally, the standard deviation increases from the 1M to the 8D product nearly by a factor of 3 (see Tables S3 and S4 in Supporting Information S1).

Finally, variability on timescales longer than 15 months (interannual mode) contributes the least to the overall standard deviation in our product (Figure 10c). The global mean temporal standard deviation amounts to 0.52 mmol m^{-2} day^{-1}, for both the 8D and 1M products. Here, the interannual variability of the equatorial Pacific driven by El-Niño—Southern Oscillation (ENSO) is the most dominant feature, but also the higher latitudes contribute substantially, with several regions having standard deviations of more than 0.8 mmol m^{-2} day^{-1}.

This attribution of the total variability of the sea-air CO_2 fluxes to three different modes of variability is qualitatively similar to that undertaken by (Gu et al., 2023). Perhaps unsurprisingly, there is good agreement between our studies with regard to the partitioning into subseasonal, seasonal, and interannual modes of variability. They found that the total energy of the seasonal variance was an order of magnitude greater than for the subseasonal and interannual modes. This aligns with our findings, where we find an even stronger relative contribution of the subseasonal mode to the total variance, particularly at the 1M resolution.

5.3. Drivers of the Variability of FCO_2

The Reynolds decomposition of the sea-air CO_2 fluxes (Equation 6) permits us to identify the main drivers for each mode of variability (Figure 11). The results of the 8D product reveal that all components contribute to the variability, but with very different magnitudes for the different timescales. Changes in ΔfCO_2 dominate the seasonal variability of FCO_2, contributing 97% to the total seasonal variance. This is also the case for the longer modes of variability, with ΔfCO_2 contributing ~60% to both the sub-decadal and decadal modes of variability. In contrast, for the subseasonal variability, the changes in ΔfCO_2 are less important, but still contribute by more than 20% to the variance.

Wind variability, which is captured in our analysis by the variations in the gas transfer coefficient, dominates the variability on subseasonal timescales, where it contributes more than 80% to the variance. Wind is also very important on sub-decadal timescales, where its contribution is more than 40%. In contrast, it matters much less on decadal and seasonal timescales. The role of the cross-term contributions is similar to that of the wind variability, with the largest contribution occurring on subseasonal and subdecadal timescales (around 40%), while its contribution to the other timescales is less than 10%.

The covariances between the different Reynolds terms (see Equation 7) are negative, thus contributing negatively to overall variability. This is a result of negative correlations between the mean state and the variability, such as

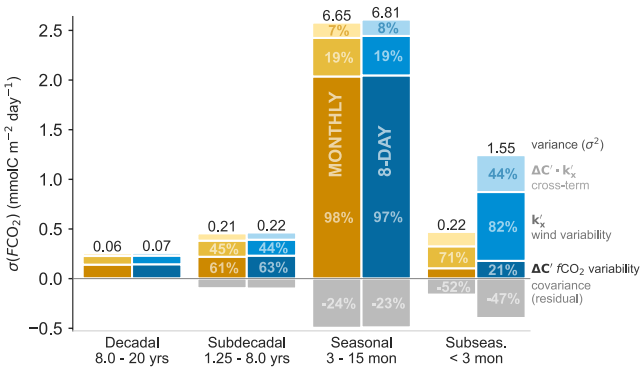

Figure 11. The contribution of $\Delta fCO_2'$ (dark shading), wind and temperature (k_x', medium shading), and the cross-term ($\Delta fCO_2' \cdot k_x'$, light shading) to each temporal mode of variability for 8-day (blue) and monthly (yellow) FCO_2. The covariances are shown in light gray. The height of the bars (positive only) shows the total standard deviation (σ) and the black numbers above each bar show the total variance (σ^2). The total height of the bar (positive and negative) represents the sum of the variances without the covariance term. The percentage contributions of k_x', $\Delta fCO_2'$, cross-term, and covariances are with respect to the total variance. Thus, multiplying σ^2 with the percentage is approximately the variance for that component (within the uncertainty of rounding). See Tables S3, S4, and S5 in Supporting Information S1 for low and high-resolution variance, standard deviation, and percentages, respectively.

occurs, for example, when regions/times with high mean wind speeds have low variations in the sea-air difference in fCO_2. This offsetting effect is particularly strong on seasonal and subseasonal timescales, where it offsets the variability by more than 20%, and nearly 50%, respectively.

Figure 11 also quantifies the strong increase in resolved subseasonal variability when moving from the 1M to the 8D product of OceanSODA-ETHZv2. The changes in variance on the other timescales are negligibly small. This indicates that as the time scales get longer, the scales of spatial variability get larger as well, so that one does not gain much in going from 1° to 0.25° resolution when one is interested in the seasonal cycle or longer timescale variability.

The more than 6-fold gain in the subseasonal variance of FCO_2 is driven by all components of the Reynolds decomposition (see also Table S5). The greatest gain comes from better-resolved wind variability as expressed in the contribution of k_x', whose σ increases nearly three times from 0.40 mmol m^{-2} day^{-1} to 1.12 mmol m^{-2} day^{-1} (1M to 8D). The σ of the cross-term variability of the subseasonal mode also increases substantially from 0.32 mmol m^{-2} day^{-1} to 0.82 mmol m^{-2} day^{-1}. This indicates an increase in the interaction between the variability of k_x and fCO_2 in the high-resolution fluxes. Mechanistically this makes sense, since wind-driven upwelling (captured by k_x) can lead to rapid changes in ΔfCO_2, thus resulting in larger FCO_2. The contribution of the ΔfCO_2 term is less important than that of the wind-driven components, but it still pushes up the variability by 0.48 mmol m^{-2} day^{-1} (Table S5 in Supporting Information S1). Particularly striking is also the very strong negative contribution of the covariance term, which reduces the overall variability by around 50%.

The strong increases in the covariance and cross-term components indicate where, perhaps, the biggest benefit is obtained by going from 1M to 8D resolution. These increases suggest that the higher resolution in time and space improves our capability of capturing the covariability of wind, surface ocean dynamics, and the ocean carbon cycle. This covariability is the result of tight process-based interactions between these drivers, providing much insight into the functioning of the ocean carbon sink (Toolsee et al., 2024). Thus, while the direct contribution of the increase in resolution in ΔfCO_2 may not seem particularly large at first sight in Figure 11, it brings a lot of knowledge gain by better capturing the processes controlling the sea-air CO_2 flux.

Our results of the Reynold decomposition for the seasonal and longer timescales are similar to those of Gu et al. (2023). The main consistency is that, on seasonal timescales, variations in k_x, that is, wind dominate, while at longer timescales, the ΔfCO_2 term dominates. More specifically, Gu et al. (2023) attributed 66% of subseasonal variability (<3 months) to the wind component, while we attribute 71% at the 1M-scale (comparable to their study).

6. Case Study: Hurricane Maria

The gain in knowledge that stems from the co-analysis of wind variations with changes in surface ocean biogeochemistry and hence surface ocean fCO_2 is best shown through an example, for which we use Hurricane Maria as a case study. Hurricane Maria occurred in September 2017 and belongs to the 10 most intense Atlantic hurricanes ever recorded. It made history as it made its first landfall in Puerto Rico before it turned northward and plowed through the northwestern North Atlantic. Hurricane Maria was most intense from the 17th to 27th of September, maintaining hurricane status throughout this period (shown by the black contours in Figures 12a–12c). Thereafter, the system moved eastward. Here, we investigate the local-scale impacts of the increase in variance by assessing the FCO_2 during the passage of Maria, which represents a short-lasting extreme event that is missed in monthly reconstructions.

At first glance, the spatial distribution of the September 2017 mean of the 8D and 1M FCO_2 look similar (Figures 12a and 12b). However, the difference between the two mean states (Figure 12c) shows that outgassing was in fact less intense along the northern part of the hurricane path for the 8D FCO_2 relative to 1M. To better

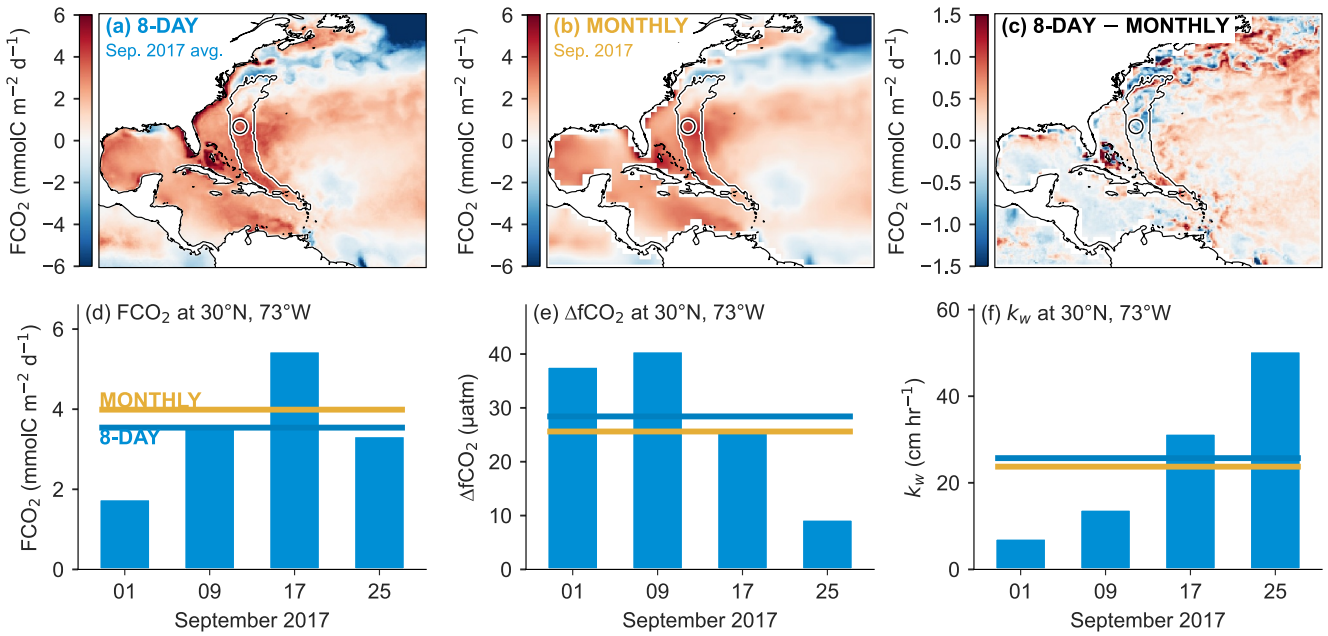

Figure 12. The mean of the 8-day (a) and monthly (by) FCO_2 for September 2017. The contour line shows the path of Hurricane Maria from 17th to 27th September, where hourly winds exceeded 20 m s−1 (c) the difference between the 8-day mean and monthly values. (d) a bar plot of FCO_2 for the four time steps in September 2017 for the location indicated by the circle in the maps (a–c). (e) and (f) represent the same, but for ΔfCO_2 and k_w respectively. The FCO_2 shown in (d) can be thought of as the product of the corresponding bars in (e) and (f) which are scaled by the solubility (not shown). Slight differences between averages (horizontal lines in e,f) are due to the marker not being at exactly the same location for high and low-resolution estimates.

understand the lower-than-expected outgassing, we plot the temporal evolution of the fluxes, ΔfCO_2 and k_w at a point location (29°N, 72°W) for each of the 8-day time steps in September 2017 (Figures 12d–12f). The location exhibits one of the strongest differences between the 8D and 1M FCO_2 for the spatio-temporal domain.

Importantly, we can also show that the response in Figure 12 is not just due to the intensification of the wind, but also due to changes in fCO_2 (Figures 12e and 12f). Our method is able to capture a decrease in ΔfCO_2 from the period centered on the seventeenth to the following period on the 25th of September, particularly between ~25°N to ~30°N (Figure 13c). The decrease (>|15| μatm) co-occurs with a reduction in sea surface temperature and an increase in chlorophyll-a (Figure 13a,b), a relationship which has been previously observed (Babin et al., 2004; Reul et al., 2021). Mechanistically, the decrease in fCO_2 is consistent with previous studies that found tropical cyclones to cool the surface ocean (decreasing fCO_2), but also to induce mixing that entrains carbon-rich waters, thus increasing fCO_2 (Yu et al., 2020). This is followed by an increase in primary productivity (i.e., a reduction in

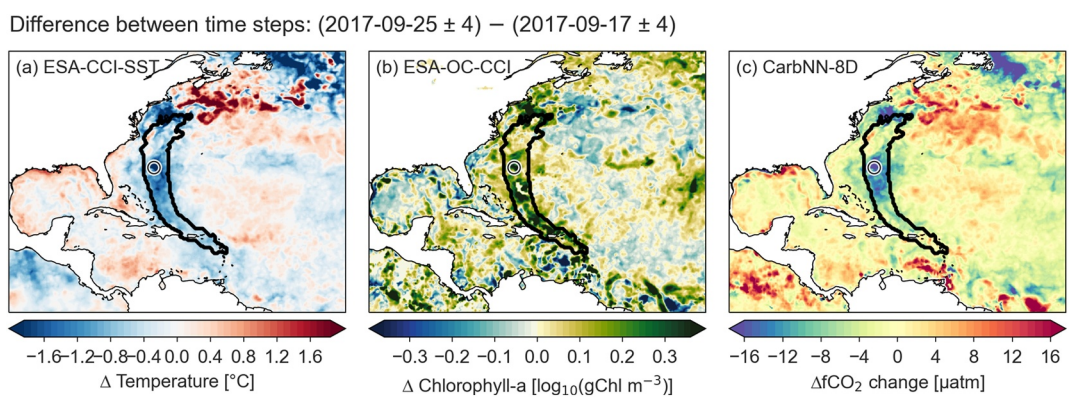

Figure 13. Maps showing the difference in (a) sea surface temperature, (b) chlorophyll-a, (c) ΔfCO_2 before and after Hurricane Maria in 2017. The two time periods are in September 13–20, and September 21–28. The black contour line shows the path of Hurricane Maria, defined by wind speeds >20 m s^{-1}.

fCO_2 Babin et al., 2004; Lévy et al., 2012; Yu et al., 2020). For our point location, SST decreased by 2.6°C between the two periods (Figure 12a), which would lead to a reduction in fCO_2 of ∼42 µatm. However, the OceanSODA-ETHZv2 product predicts a fCO_2 reduction of 24 µatm, leaving an excess of 18 µatm, which we attribute to the entrainment of DIC-rich waters. However, an increase in chlorophyll-a would result in a further decrease in fCO_2, meaning that the contribution of entrainment could be even larger; however, this contribution cannot be empirically determined.

Although OceanSODA-ETHZv2 captures the impacts of Hurricane Maria on FCO_2 and fCO_2, the magnitude of the event is probably underestimated. For example, observation-based studies found that high-velocity winds increased outgassing by >30 mmol m^{-2} day^{-1} for ∼24 hr (Ye et al., 2020; Yu et al., 2020), compared to the 2 mmol m^{-2} day^{-1} increase observed from 9 to 17 September (Figure 12d). The 8D resolution averages the short-lasting spikes (∼24 hr). Furthermore, satellite and reanalysis products may underestimate the spikes; for example, ERA5 underestimates extreme wind conditions by between 5% and 10% (relative to satellites Campos et al., 2022). Future work could address this by investigating the influence of using 8-day ΔfCO_2 with hourly, daily, 8-day and monthly k_w using an extreme wind speed specific wind speed data set (e.g., https://www.maxss.org).

7. Discussion

We provide here for the first time a global-scale gap filled fCO_2 product at an 8 days and 0.25° × 0.25° resolution, from which we can compute the sea-air CO_2 fluxes at the same unprecedented resolution. But what are we gaining from this increase in resolution in terms of quality and what are we learning from this in terms of processes? Next, we discuss these two questions in turn.

7.1. Impact of High Resolution on fCO_2 Mapping Uncertainty

In their work investigating a range of gap filled fCO_2 products at 1M resolution, Gregor et al. (2019) found that all products had very similar RMSD ranging between 15 and 20 µatm. They proposed that this is a methodological wall beyond which it would be difficult to progress. Our question is thus, can higher resolution get us over the wall? In the open ocean, we find little reduction in the uncertainties, that is, the RMSD of our 8D product in the open ocean (∼13 µatm) is similar to that of previous approaches with 1M resolution (Figure 3f). In addition, relatively large biases still occur in some poorly sampled regions (Figures 5 and 6). In the coastal ocean there is some gain, with OceanSODA-ETHZv2 having a lower RMSD (25.4 µatm) relative to its predecessor (Ocean-SODA-ETHZ = 27 µatm), a monthly by 1° pCO_2-product (Gregor & Gruber, 2021). An improvement in coastal ocean RMSD was also reported by Chau et al. (2024) in the CMEMS-FFNN approach (28.5 µatm to 27.6 µatm) by increasing the spatial resolution from 1° to 0.25° (though not the temporal resolution). The slight reduction in coastal RMSD in both approaches is probably due to the fact that sharper gradients and fine-scale features are better represented in high-resolution estimates (e.g., Figure 8). In other words, there is a better match-up between the gridded fCO_2 observations and the gridded predictors, also called the representation uncertainty by (Gregor & Gruber, 2021). However, coastal uncertainties remain large and the 8D product struggles to fully capture near-shore dynamics (Tables S3 and S4 in Supporting Information S1).

Thus, while we were able to reduce the random uncertainties of the gap-filled fCO_2 product at 8D resolution, this is not a breakthrough. Some further reduction may be achieved by going to even higher resolution, especially in time, as this would permit one to capture the more ephemeral or faster developing features, such as the tropical instability waves and short-lived upwelling events. The big challenge here is the global-scale availability of predictor variables. The continuing development of global ocean re-analyses that also increasingly incorporate ocean color (e.g., Green Mercator) or commercial satellite observations at daily and 10 m spatial resolutions (Shutler et al., 2024) may soon provide such very high-frequency predictors.

At the same time, we may be at the limit of the capability of the current generation of gap-filling methods that use simple architectures. With the current architecture, where each sample is independent in space and time, these gap-filling methods will struggle to solve two problems: (a) the basin-scale problem, where high accuracy of the inferred fCO_2 is absolutely essential for constraining the ocean carbon sink, and (b) a fine-scale problem, where drivers can cause rapid and short spatial-scale variations in fCO_2 that need to be tracked with high precision. A potential solution to this problem is to explore the use of machine learning methods that retain information at the basin scale (e.g., exploit temporal evolution or correlation), while considering patterns of local variability (e.g.,

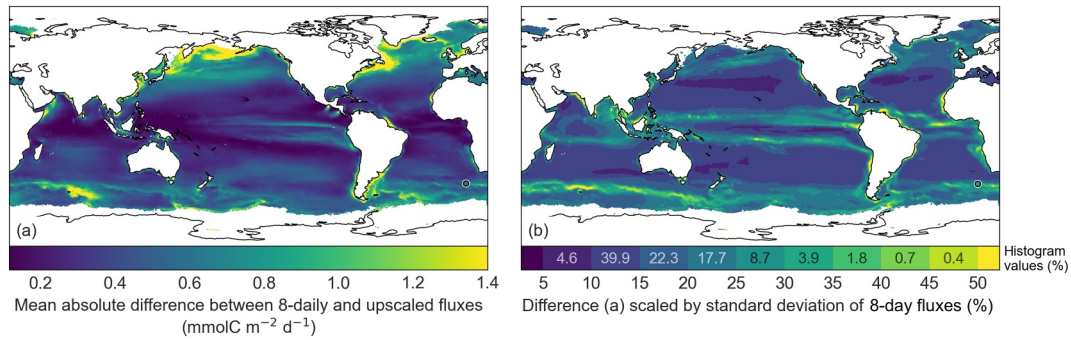

Figure 14. (a) The mean absolute difference (MAD) between FCO_2^{8D} and $FCO_2^{upscaled}$, where the latter is calculated with k_x^{8D} and ΔfCO_2^{IM} upscaled to 8D. Thus, (a) shows the impact of using high-resolution ΔfCO_2, even if a high-resolution gas transfer velocity is used. (b) shows the MAD from (a) scaled to the standard deviation of the high-resolution fluxes (Figure 9b). The values in the color bar represent the percentage cover that the histogram bin occupies in (b).

filaments). Machine learning and computer vision techniques that exploit spatio-temporal features, such as those used in weather-forecasting neural networks (e.g., GraphCast and FourCastNet; Lam et al., 2023; Pathak et al., 2022) and image-based generative neural networks (Rombach et al., 2022) may offer solutions to this multiscale problem.

7.2. Lessons From High-Resolution FCO_2

Our results show that there are distinct benefits of high-resolution predictions of fCO_2 and sea-air CO_2 flux. First, the high-resolution predictions of fCO_2 reduce the uncertainty of the estimates relative to past low-resolution estimates in the coastal ocean due to improved match-ups between the in-situ target data (i.e., SOCAT) and the remote predictors (Section 7.1). In addition, higher resolution captures more variability of fCO_2 and sea-air CO_2 fluxes (Figure 11). However, we also show that the majority of the gained variability, particularly at the subseasonal scale, is due to an increase in the temporal resolution of wind represented by k_x. Still, we also showed that the higher resolution fCO_2 product permits us to better capture the covariability of fCO_2 and wind, and hence the interaction between physical forcing and the ocean carbon cycle as shown by in-situ observations (Toolsee et al., 2024).

Since winds drive the majority of the increase in variability, a key question is whether a similar result can be achieved by "upscaling" the fluxes. The upscaled fluxes $\left(FCO_2^{upscaled}\right)$ are calculated using low-resolution fCO_2 and high-resolution k_x. The difference between FCO_2^{8D} and $FCO_2^{upscaled}$ (Figure 14a) quantifies the amount of information missed when relying solely on $FCO_2^{upscaled}$. In large parts of the ocean, particularly the gyre regions, we find minimal differences between FCO_2^{8D} and $FCO_2^{upscaled}$ (<0.4 mmol m^{-2} d^{-1}, as shown in Figure 14a). However, in dynamic regions of the ocean, the differences can be substantial (>1 mmol m^{-2} d^{-1}).

Scaling these results by the standard deviation of FCO_2^{8D} indicates the local importance of using the high-resolution FCO_2 (Figure 14b). The Antarctic Circumpolar Current (ACC) is the most striking feature, where significant additional variability is resolved as a result of the high-resolution. Interestingly, this aligns with an observation-based study in the Atlantic sector of the ACC that suggested a 3-daily sampling frequency of fCO_2 is required to constrain fluxes to a 10% uncertainty threshold (location indicated by the marker in Figure 14; Monteiro et al., 2015). This also holds true for the eastern boundary upwelling regions, and the continental shelf regions in the high-latitude Pacific and Atlantic. Thus, it is in regions where there is the combination of high spatio-temporal variability and high wind speeds that drives the large differences between FCO_2^{8D} and $FCO_2^{upscaled}$. However, it is also important to recognize that more sporadic events, such as Hurricane Maria, are also captured, but they may contribute little to the large-scale variability due to their short-lived nature (Lévy et al., 2012). While a step in the right direction, we also have to note that the magnitude of the variability is probably underestimated at the 8-day resolution, for example, Hurricane Maria (Lévy et al., 2012; Ye et al., 2020; Yu et al., 2020).

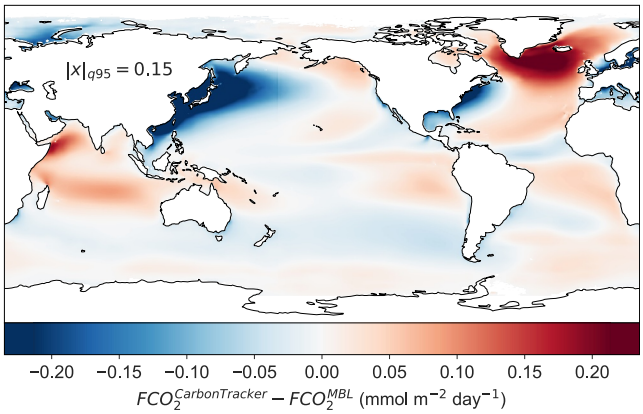

Figure 15. Map of the difference between sea-air CO_2 fluxes computed with xCO_2 from NOAA's CarbonTracker v2022 (CT2022) and Marine Boundary Layer (MBL) products. Positive (red) shows regions of increased outgassing (or decreased uptake) and negative values show regions of increased uptake (or decreased outgassing) for the CarbonTracker fluxes. The value, $|x|_{q95}$ shows the 95 percentile of the absolute value of the differences shown in the figure.

Despite the increases in local-scale variability, we do not see major differences between the mean sea-air CO_2 fluxes at 8-day and monthly resolutions at global and interannual scales (Figures 9a and 11). Thus, if the aim is to constrain FCO_2 at large scales, for example, for the Global Carbon Budget (Friedlingstein et al., 2022), our analysis suggests that there is little reason to use 8-day by 0.25° estimates. Importantly, increasing the resolution of FCO_2 does not solve current unanswered problems and questions. For example, it does not close the growing divide in the global CO_2 uptake between the fCO_2-products and Global Ocean Biogeochemical Models (GOBMs) over the last decade (2010−2022) (Hauck et al., 2023). Several studies have suggested that this divergence is a result of the overestimation of decadal variability by the fCO_2-products (Gloege et al., 2021; Hauck et al., 2023). However, the evidence is not conclusive and it may well be that GOBMs underestimate decadal variability (Mayot et al., 2023). However, the fact remains that fCO_2-products suffer from observational paucity and sampling biases (Gloege et al., 2021; Hauck et al., 2023; Ritter et al., 2017; Rödenbeck et al., 2015), a problem that the high-resolution FCO_2 estimates presented here cannot solve.

8. Caveats

There are a number of specific caveats and challenges that need to be considered in our novel 8D product. The first concerns the use of the NOAA marine boundary layer product (MBL) for the atmospheric dry air mixing ratio $\left(xCO_2^{\text{mbl}}\right)$ for the computation of the sea-air CO_2 difference. The second concerns the use of a stacked salinity product in order to produce a high-resolution product spanning four decades.

While the NOAA xCO_2^{mbl} product resolves the dominant mode of spatial variations (i.e., the latitudinal gradient) in the atmospheric dry air-mixing ratio (Tans & Keeling, 2023), it fails to resolve any zonal variations. Palter et al. (2023) pointed out recently that this neglect may lead to significant uncertainties in the sea-air CO_2 fluxes, especially downwind of the major continents where most of the fossil fuel CO_2 is being emitted. To investigate the potential impact of our choice to use the MBL product, we computed the sea-air CO_2 fluxes instead with the zonally resolved NOAA CarbonTracker product for atmospheric xCO_2^{CT2022}, version CT2022 (Jacobson et al., 2023). We used exactly the same procedures to compute the atmospheric fCO_2 from the dry air mixing ratio as we used for the MBL product.

The use of CarbonTracker for atmospheric CO_2 has little to no impact on the globally integrated sea-air CO_2 flux. However, this is not true for its spatial distribution, particularly for regions downwind of industrialized countries with high anthropogenic CO_2 emissions (Figure 15), confirming the results of Palter et al. (2023). The magnitude of these differences can be up to 5% of the total variability for coastal oceans in the lee of continental emissions. The comparison also reveals substantial reductions in ocean uptake in the North Atlantic and in the Indian Ocean, driven by lower atmospheric fCO_2 in the CarbonTracker product. The reduction in the North Atlantic is likely a consequence of the strong ocean uptake flux in this region, which tends to lower atmospheric fCO_2, and effect not captured in the zonally averaged MBL product.

Thus, the use of xCO_2^{CT2022} instead of xCO_2^{MBL} in the calculation of FCO_2 can be regionally significant. However, the current CarbonTracker (CT2022) product is available only from 2000 onward (Jacobson et al., 2023). Thus, for this analysis, we continue to use the MBL product. This choice also allows us to remain consistent with nearly all previous work on ocean mapping of surface ocean fCO_2. However, in the numerical data set, we provide both sea-air CO_2 flux estimates for the post-2000 period. However, we recommend further exploration of the benefits of using spatially better-resolved atmospheric CO_2 products, especially when exploring coastal and/or high-resolution fluxes.

The second caveat concerns the difficulties associated with the estimation of fCO_2 over a period of four decades, given the limitations in the availability of the predictors of choice (Figure 2). For chlorophyll and sea surface height, this means that climatologies are used for the periods where there is no coverage (prior to 1998 and 1993

respectively). Given that climatologies have smoother fields, the fCO_2 estimates reflect these smoother input data. In the case of salinity, three different products are used during the four-decade period (Figure 2 and Figure S5a in Supporting Information S1).

The difference in variability between salinity products can be large for some regions (e.g., Figure S5e in Supporting Information S1). For most of the open ocean, the difference between the salinity products does not make a significant difference, as salinity is a weak driver of ΔfCO_2 (<|5| µatm PSU^{-1}; Figure S5a in Supporting Information S1). However, in the equatorial Pacific, particularly in the western part of the basin, the sensitivity of ΔfCO_2 to salinity is large (>20 µatm PSU^{-1}). This means that salinity anomalies typically drive a change of more than ∼10 µatm in the western equatorial Pacific for the SODA v3.4.2 salinity (Figure S5b). For the ESA-CCI salinity, the salinity climate data record product, salinity drives an average change of ∼6 µatm.

We stress that this only affects the variability of ΔfCO_2 since the mean state is captured by the 8-day climatology of salinity. However, it does mean that variability of fCO_2 in the western equatorial Pacific, may be overestimated in the period 1982–1992 (SODA salinity) and underestimated from 1993 to 2009 (CMEMS-Multiobs). This applies particularly to the large-scale changes in salinity driven by El Niño. While this is a lesser problem for ΔfCO_2 predictions, this will have a much stronger influence on machine learning estimates of total alkalinity (e.g., Gregor & Gruber, 2021)

9. Conclusions and Next Steps

In this study, we present the first 8-day by $0.25° \times 0.25°$ estimates of sea-air CO_2 fluxes at a global scale over four decades. The high-resolution OceanSODA-ETHZv2 fCO_2-product is able to capture significantly more variability in sea-air CO_2 flux, with the majority of the increased variability being driven by higher resolution winds, rather than fCO_2. However, we also show that high-resolution fCO_2 is important in regions with high variability. This includes short-lived high-intensity events, such as upwelling events and hurricanes. In particular, we also demonstrate that the higher resolution fCO_2 product allows us to better analyze the interaction between variations in wind and the ocean carbon cycle, thus improving our understanding of the processes.

Following the approach of (Gregor & Gruber, 2021), the fCO_2 estimates from the OceanSODA-ETHZv2 approach can be combined with high-resolution estimates of total alkalinity to estimate high-resolution ocean acidification parameters that could be used to better understand ocean acidification extremes, for example, (Burger et al., 2020; Desmet et al., 2022; Gruber et al., 2021). Understanding such extreme events from historical data is important so that the drivers of extremes can be better characterized (Gruber et al., 2021). This then also allows us to understand current extreme events, such as the North Atlantic marine heatwave in 2023, that could drive anomalous changes in ocean acidification. This requires near-real-time capability of machine learning approaches, which is technically quite feasible. However, the current release cycle of SOCAT means that near-real time estimates would be predicting up to 1.5 years beyond the target data (D. C. Bakker et al., 2016). The impact of predicting beyond the target data needs to be investigated before results can be used.

Another area of improvement lies in the temporal resolution of our sea-air CO_2 fluxes. The psuedo-daily estimates of sea-air CO_2 fluxes could quite simply be estimated by upscaling ΔfCO_2 to daily resolution while using daily estimates of k_w. This could be further improved by using daily predictors for those that are available (e.g., SST, and SSH) alongside upscaled predictors for those that are not. However, the gain from 8-day to daily fCO_2 at the current spatial resolution (0.25°) is unlikely to be as large as the improvement from monthly to 8-day (Monteiro et al., 2015).

We also need to improve the way we consider variations in atmospheric CO_2 for the estimation of the sea-air CO_2 fluxes. While the NOAA MBL product has served the community well, further improvement of our ability to quantify the variations of the ocean carbon sink requires us to move to a product that resolves the meridional and zonal variations in atmospheric CO_2. The NOAA CarbonTracker product is a good candidate, but by covering only the period from 2000 onward, its use is currently limited. Extending the NOAA CarbonTracker back in time is highly desirable. Other products may be available and shall be explored as well.

Finally, the field of machine learning is developing at an unprecedented rate. New approaches such as Fourier neural operators used in Nvidia's FourcastNet could be incorporated to better capture fine-scale variability of fCO_2. However, it is unclear if even these approaches would be able to reduce the uncertainties beyond "the wall" that sits between 15 and 20 µatm.

Data Availability Statement

The OceanSODA-ETHZv2 data set of ΔfCO_2 and FCO_2 produced and used throughout this study (Gregor, Gruber, & Shutler, 2024) are available at https://doi.org/10.5281/zenodo.11206366. Code to create the data and figures (Gregor, Shutler, & Gruber, 2024) is hosted at https://doi.org/10.5281/zenodo.11230362. All data used to create the above-mentioned data set are at least open-access under academic license and are listed here. SOCAT v2023 data was downloaded from https://socat.info/socat_files/v2023/SOCATv2023.tsv.zip (D. C. Bakker et al., 2016; D. C. E. Bakker et al., 2023). Sea-surface temperature is from https://doi.org/10.48670/moi-00169 (Good et al., 2020). ERA5 data (wind and sea level pressure) are from https://doi.org/10.24381/cds.adbb2d47 (Hersbach et al., 2020, 2023). Salinity from 2010 to 2020 is from https://catalogue.ceda.ac.uk/uuid/fad2e982a59d44788eda09e3c67ed7d5 (Boutin et al., 2021). Salinity and mixed layer depth from SODA v3.4.2 were downloaded from https://dsrs.atmos.umd.edu/DATA/soda3.4.2/REGRIDED/ocean/ (Carton et al., 2018). Salinity after 2021 was downloaded from https://doi.org/10.48670/moi-00051 (Droghei et al., 2016). Chlorophyll-a data can be found at https://www.oceancolour.org/ (Sathyendranath et al., 2023). We used the reprocessed sea surface height from https://doi.org/10.48670/moi-00148 (see acknowledgments).

Acknowledgments

We are indebted to the large number of scientists who have measured the surface ocean fCO_2 over the past few decades and to the programs and agencies that enabled or funded these observations. We are also very grateful to all those who have assembled and quality-controlled these observations into readily usable products. A special thanks to the Surface Ocean CO2 Atlas (SOCAT) team under the leadership of Dorothee Bakker. SOCAT is an international effort, endorsed by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS) and the Integrated Marine Biosphere Research (IMBeR) program, to deliver a uniformly quality-controlled surface ocean CO_2 database. We are also grateful to many other programs that have collected and processed marine observations in the last few decades, enabling the development of novel statistical techniques for interpreting the CO_2 observations. The work of LG, JS, and NG was supported by the ESA OceanHealth-OA project (contract number 4000137603/22/I-DT). JS has received further support by the ESA MAXSS project (4000132954/20/I-NB). The Ssalto/Duacs altimeter products were produced and distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS) (http://www.marine.copernicus.eu). We thank Rik Wanninkhof and two additional reviewers for their thorough reviews and their great comments that helped to improve this study. Open access funding provided by Eidgenossische Technische Hochschule Zurich.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. [software]. *Tensorflow*. Retrieved from https://www.tensorflow.org/

Alvera-Azcárate, A., Barth, A., Sirjacobs, D., Lenartz, F., & Beckers, J. M. (2011). Data interpolating empirical orthogonal functions (DINEOF): A tool for geophysical data analyses. *Mediterranean Marine Science*, *12*(3), 5–11. https://doi.org/10.12681/mms.64

Arruda, R., Calil, P. H. R., Bianchi, A. A., Doney, S. C., Gruber, N., Lima, I., & Turi, G. (2015). Air-sea CO_2 fluxes and the controls on ocean surface pCO_2 seasonal variability in the coastal and open-ocean southwestern Atlantic Ocean: A modeling study. *Biogeosciences*, *12*(19), 5793–5809. https://doi.org/10.5194/bg-12-5793-2015

Babin, S. M., Carton, J. A., Dickey, T. D., & Wiggert, J. D. (2004). Satellite evidence of hurricane-induced phytoplankton blooms in an oceanic desert. *Journal of Geophysical Research*, *109*(C3). https://doi.org/10.1029/2003JC001938

Bakker, D. C., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., et al. (2016). A multi-decade record of high-quality fCO2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT). *Earth System Science Data*, *8*(2), 383–413. https://doi.org/10.5194/essd-8-383-2016

Bakker, D. C. E., Alin, S. R., Bates, N., Becker, M., Feely, R. A., Gkritzalis, T., et al. (2023). Surface ocean CO_2 atlas database version 2023 (SOCATv2023) (NCEI accession 0278913). [Dataset]. *NOAA National Centers for Environmental Information*. https://doi.org/10.25921/r7xa-bt92

Bates, N. R., Astor, Y. M., Church, M. J., Currie, K., Dore, J. E., González-Dávila, M., et al. (2014). A time-series view of changing surface ocean chemistry due to ocean uptake of anthropogenic CO_2 and ocean acidification. *Oceanography*, *27*(1), 126–141. https://doi.org/10.5670/oceanog.2014.16

Bates, N. R., Takahashi, T., Chipman, D. W., & Knap, A. H. (1998). Variability of pCO_2 on diel to seasonal timescales in the Sargasso Sea near Bermuda. *Journal of Geophysical Research*, *103*(C8), 15567–15585. https://doi.org/10.1029/98JC00247

Bennington, V., Galjanic, T., & McKinley, G. A. (2022). Explicit physical knowledge in machine learning for Ocean Carbon flux reconstruction: The pCO_2-residual method. *Journal of Advances in Modeling Earth Systems*, *14*(10). https://doi.org/10.1029/2021MS002960

Boutin, J., Merlivat, L., HÉnocq, C., Martin, N., & SallÉe, J. B. (2008). Air-sea CO_2 flux variability in frontal regions of the Southern Ocean from CARbon Interface OCean Atmosphere drifters. *Limnology & Oceanography*, *53*(5part2), 2062–2079. https://doi.org/10.4319/lo.2008.53.5_part_2.2062

Boutin, J., Vergely, J. L., Marchand, S., D'Amico, F., Hasson, A., Kolodziejczyk, N., et al. (2018). New SMOS Sea Surface Salinity with reduced systematic errors and improved variability. *Remote Sensing of Environment*, *214*(August 2017), 115–134. https://doi.org/10.1016/j.rse.2018.05.022

Boutin, J., Vergely, J.-L., Reul, N., Catany, R., Koehler, J., Martin, A., et al. (2021). ESA sea surface salinity climate change initiative (Sea_-Surface_Salinity_cci): Weekly sea surface salinity product, v03.21, for 2010 to 2020. [Dataset]. *NERC EDS Centre for Environmental Data Analysis*. https://catalogue.ceda.ac.uk/uuid/fad2e982a59d44788eda09e3c67ed7d5

Burger, F. A., John, J. G., & Frölicher, T. L. (2020). Increase in ocean acidity variability and extremes under increasing atmospheric CO2. *Biogeosciences*, *17*(18), 4633–4662. https://doi.org/10.5194/bg-17-4633-2020

Burt, W. J., Thomas, H., Miller, L. A., Granskog, M. A., Papakyriakou, T. N., & Pengelly, L. (2016). Inorganic carbon cycling and biogeochemical processes in an Arctic inland sea (Hudson Bay). *Biogeosciences*, *13*(16), 4659–4671. https://doi.org/10.5194/BG-13-4659-2016

Campos, R. M., Gramcianinov, C. B., de Camargo, R., & da Silva Dias, P. L. (2022). Assessment and calibration of ERA5 severe winds in the Atlantic Ocean using satellite data. *Remote Sensing*, *14*(19), 4918. https://doi.org/10.3390/rs14194918

Carton, J. A., Chepurin, G. A., & Chen, L. (2018). SODA3: A new ocean climate reanalysis. *Journal of Climate*, *31*(17), 6967–6983. https://doi.org/10.1175/JCLI-D-17-0149.1

Chapa-Balcorta, C., Hernandez-Ayon, J. M., Durazo, R., Beier, E., Alin, S. R., & López-Pérez, A. (2015). Influence of post-Tehuano oceanographic processes in the dynamics of the CO_2 system in the Gulf of Tehuantepec, Mexico. *Journal of Geophysical Research: Oceans*, *120*(12), 7752–7770. https://doi.org/10.1002/2015JC011249

Chau, T. T. T., Gehlen, M., & Chevallier, F. (2022). A seamless ensemble-based reconstruction of surface ocean pCO_2 and air-sea CO_2 fluxes over the global coastal and open oceans. *Biogeosciences*, *19*(4), 1087–1109. https://doi.org/10.5194/bg-19-1087-2022

Chau, T.-T.-T., Gehlen, M., Metzl, N., & Chevallier, F. (2024). CMEMS-LSCE: A global, 0.25°, monthly reconstruction of the surface ocean carbonate system. *Earth System Science Data*, *16*(1), 121–160. https://doi.org/10.5194/essd-16-121-2024

Chen, S., Hu, C., Barnes, B. B., Wanninkhof, R., Cai, W.-J., Barbero, L., & Pierrot, D. (2019). A machine learning approach to estimate surface ocean pCO_2 from satellite measurements. *Remote Sensing of Environment*, *228*, 203–226. https://doi.org/10.1016/j.rse.2019.04.019

Courtois, A., Morel, J. M., & Arias, P. (2023). Can neural networks extrapolate? Discussion of a theorem by Pedro Domingos. Revista de la Real Academia de Ciencias Exactas. *Fisicas y Naturales - Serie A: Matematicas*, *117*(2), 1–26. https://doi.org/10.1007/s13398-023-01411-z

Denvil-Sommer, A., Gehlen, M., Vrac, M., & Mejia, C. (2019). LSCE-FFNN-v1: A two-step neural network model for the reconstruction of surface ocean pCO_2 over the global ocean. *Geoscientific Model Development*, *12*(5), 2091–2105. https://doi.org/10.5194/gmd-12-2091-2019

Desmet, F., Gruber, N., Köhn, E. E., Münnich, M., & Vogt, M. (2022). Tracking the space-time evolution of ocean acidification extremes in the California current system and Northeast Pacific. *Journal of Geophysical Research: Oceans*, *127*(5), e2021JC018159. https://doi.org/10.1029/2021JC018159

DeVries, T., Le Quéré, C., Andrews, O., Berthet, S., Hauck, J., Ilyina, T., et al. (2019). Decadal trends in the ocean carbon sink. *Proceedings of the National Academy of Sciences*, *116*(24), 201900371–201911651. https://doi.org/10.1073/pnas.1900371116

DeVries, T., Yamamoto, K., Wanninkhof, R., Gruber, N., Hauck, J., Müller, J. D., et al. (2023). Magnitude, trends, and variability of the global ocean carbon sink from 1985-2018. *Global Biogeochemical Cycles*, *37*(10), e2023GB007780. https://doi.org/10.1029/2023GB007780

Doney, S. C., Lima, I., Feely, R. A., Glover, D. M., Lindsay, K., Mahowald, N., et al. (2009). Mechanisms governing interannual variability in upper-ocean inorganic carbon system and air–sea CO2 fluxes: Physical climate and atmospheric dust. *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(8–10), 640–655. https://doi.org/10.1016/j.dsr2.2008.12.006

Droghei, R., Buongiorno Nardelli, B., & Santoleri, R. (2016). Combining in situ and satellite observations to retrieve salinity and density at the ocean surface. *Journal of Atmospheric and Oceanic Technology*, *33*(6), 1211–1223. https://doi.org/10.1175/JTECH-D-15-0194.1

Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M., et al. (2021). SeaFlux: Harmonization of air–sea CO_2 fluxes from surface pCO_2 data products using a standardized approach. *Earth System Science Data*, *13*(10), 4693–4710. https://doi.org/10.5194/essd-13-4693-2021

Fay, A. R., & McKinley, G. A. (2014). Global open-ocean biomes: Mean and temporal variability. *Earth System Science Data*, *6*(2), 273–284. https://doi.org/10.5194/essd-6-273-2014

Fiedler, P. C. (2002). The annual cycle and biological effects of the Costa Rica Dome. *Deep Sea Research Part I: Oceanographic Research Papers*, *49*(2), 321–338. https://doi.org/10.1016/S0967-0637(01)00057-7

Friederich, G. E., Ledesma, J., Ulloa, O., & Chavez, F. P. (2008). Air–sea carbon dioxide fluxes in the coastal southeastern tropical Pacific. *Progress in Oceanography*, *79*(2), 156–166. https://doi.org/10.1016/j.pocean.2008.10.001

Friederich, G. E., Walz, P. M., Burczynski, M. G., & Chavez, F. P. (2002). Inorganic carbon in the central California upwelling system during the 1997-1999 El Niño-La Niña event. *Progress in Oceanography*, *54*(1–4), 185–203. https://doi.org/10.1016/S0079-6611(02)00049-6

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., et al. (2023). Global carbon Budget 2023. *Earth System Science Data*, *15*(12), 5301–5369. https://doi.org/10.5194/essd-15-5301-2023

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., et al. (2022). Global carbon budget 2022. *Earth System Science Data*, *14*(11), 4811–4900. https://doi.org/10.5194/essd-14-4811-2022

Fukushima, K. (1969). Visual feature extraction by a multilayered network of analog threshold elements. *IEEE Transactions on Systems Science and Cybernetics*, *5*(4), 322–333. https://doi.org/10.1109/TSSC.1969.300225

Gade, K. (2010). A non-singular horizontal position representation. *Journal of Navigation*, *63*(03), 395–417. https://doi.org/10.1017/S0373463309990415

Gloege, L., McKinley, G. A., Landschützer, P., Fay, A. R., Frölicher, T. L., Fyfe, J. C., et al. (2021). Quantifying errors in observationally based estimates of ocean carbon sink variability. *Global Biogeochemical Cycles*, *35*(4), 1–14. https://doi.org/10.1029/2020GB006788

Gloege, L., Yan, M., Zheng, T., & McKinley, G. A. (2022). Improved quantification of ocean carbon uptake by using machine learning to merge global models and pCO_2 data. *Journal of Advances in Modeling Earth Systems*, *14*(2), 1–19. https://doi.org/10.1029/2021MS002620

González, M. F., & Ilyina, T. (2016). Impacts of artificial ocean alkalinization on the carbon cycle and climate in Earth system simulations. *Geophysical Research Letters*, *43*(12), 6493–6502. https://doi.org/10.1002/2016GL068576

Good, S. A., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., et al. (2020). The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sensing*, *12*(4), 1–20. https://doi.org/10.3390/rs12040720

Goyet, C., Bradshaw, A. L., & Brewer, P. G. (1991). The carbonate system in the Black Sea. *Deep-Sea Research, Part A: Oceanographic Research Papers*, *38*(Suppl. 2A), S1049–S1068. https://doi.org/10.1016/S0198-0149(10)80023-8

Gregor, L., & Gruber, N. (2021). OceanSODA-ETHZ: A global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of Ocean acidification. *Earth System Science Data*, *13*(2), 777–808. https://doi.org/10.5194/essd-13-777-2021

Gregor, L., Gruber, N. P., & Shutler, J. (2024). Code for "High resolution variability of the ocean carbon sink" (v1.0). [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.11230362

Gregor, L., Lebehot, A. D., Kok, S., & Scheel Monteiro, P. M. (2019). A comparative assessment of the uncertainties of global surface ocean CO_2 estimates using a machine-learning ensemble (CSIR-ML6 version 2019a) - Have we hit the wall? *Geoscientific Model Development*, *12*(12), 5113–5136. https://doi.org/10.5194/gmd-12-5113-2019

Gregor, L., Shutler, J., & Gruber, N. (2024). OceanSODA-ETHZ-v2: Surface ocean sea-air CO_2 fluxes from 1982 to 2022 (8-day by 0.25° x 0.25°) (v2.2024r01). [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.11206366

Gruber, N., Bakker, D. C. E., DeVries, T., Gregor, L., Hauck, J., Landschützer, P., et al. (2023). Trends and variability in the ocean carbon sink. *Nature Reviews Earth & Environment*, *4*(2), 119–134. https://doi.org/10.1038/s43017-022-00381-x

Gruber, N., Boyd, P. W., Frölicher, T. L., & Vogt, M. (2021). Biogeochemical extremes and compound events in the ocean. *Nature*, *600*(7889), 395–407. https://doi.org/10.1038/s41586-021-03981-7

Gruber, N., Clement, D., Carter, B. R., Feely, R. A., van Heuven, S., Hoppema, M., et al. (2019). The oceanic sink for anthropogenic CO_2 from 1994 to 2007. *Science*, *363*(6432), 1193–1199. https://doi.org/10.1126/science.aau5153

Gruber, N., Landschützer, P., & Lovenduski, N. S. (2019). The variable southern ocean carbon sink. *Annual Review of Marine Science*, *11*(1), 159–186. https://doi.org/10.1146/annurev-marine-121916-063407

Gu, Y., Katul, G. G., & Cassar, N. (2023). Multiscale temporal variability of the global air-sea CO_2 flux anomaly. *Journal of Geophysical Research: Biogeosciences*, *128*(6). https://doi.org/10.1029/2022JG006934

Hauck, J., Nissen, C., Landschützer, P., Rödenbeck, C., Bushinsky, S., & Olsen, A. (2023). Sparse observations induce large biases in estimates of the global ocean CO_2 sink: An ocean model subsampling experiment. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *381*(2249). https://doi.org/10.1098/rsta.2022.0063

Hauck, J., Zeising, M., Le Quéré, C., Gruber, N., Bakker, D. C. E., Bopp, L., et al. (2020). Consistency and challenges in the ocean carbon sink estimate for the global carbon budget. *Frontiers in Marine Science*, *7*, 571720. https://doi.org/10.3389/fmars.2020.571720

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2023). Era5 hourly data on single levels from 1940 to present. [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.adbb2d47

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Holder, C., & Gnanadesikan, A. (2021). Can machine learning extract the mechanisms controlling phytoplankton growth from large-scale observations? A proof-of-concept study. *Biogeosciences*, *18*(6), 1941–1970. https://doi.org/10.5194/BG-18-1941-2021

Iida, Y., Takatani, Y., Kojima, A., & Ishii, M. (2021). Global trends of ocean CO_2 sink and ocean acidification: An observation-based reconstruction of surface ocean inorganic carbon variables. *Journal of Oceanography*, *77*(2), 323–358. https://doi.org/10.1007/s10872-020-00571-5

Jacobson, A. R., Schuldt, K. N., Tans, P., Andrews, A., Miller, J. B., Oda, T., et al. (2023). CarbonTracker CT2022. [dataset]. *NOAA Global Monitoring Laboratory*. https://doi.org/10.25925/Z1GJ-3254

Jähne, B., Heinz, G., & Dietrich, W. (1987). Measurement of the diffusion coefficients of sparingly soluble gases in water. *Journal of Geophysical Research*, *92*(C10), 10767–10776. https://doi.org/10.1029/JC092iC10p10767

Jones, S. D., Le Quéré, C., & Rdenbeck, C. (2012). Autocorrelation characteristics of surface ocean pCO_2 and air-sea CO_2 fluxes. *Global Biogeochemical Cycles*, *26*(2), 1–12. https://doi.org/10.1029/2010GB004017

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 3149–3157). Curran Associates Inc.

Khatiwala, S., Tanhua, T., Mikaloff Fletcher, S. E., Gerber, M., Doney, S. C., Graven, H. D., et al. (2013). Global ocean storage of anthropogenic carbon. *Biogeosciences*, *10*(4), 2169–2191. https://doi.org/10.5194/bg-10-2169-2013

Koch, J., McKinley, G. A., Bennington, V., & Ullman, D. (2009). Do hurricanes cause significant interannual variability in the air-sea CO_2 flux of the subtropical North Atlantic? *Geophysical Research Letters*, *36*(7). https://doi.org/10.1029/2009GL037553

Lachkar, Z., & Gruber, N. (2013). Response of biological production and air-sea CO_2 fluxes to upwelling intensification in the California and Canary Current Systems. *Journal of Marine Systems*, *109–110*, 149–160. https://doi.org/10.1016/j.jmarsys.2012.04.003

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416–1421. https://doi.org/10.1126/science.adi2336

Lan, X., Tans, P., Thoning, K., & NOAA Global Monitoring Laboratory. (2023). NOAA Greenhouse Gas Marine Boundary Layer Reference - CO_2 [Dataset]. *NOAA GML*. https://doi.org/10.15138/DVNP-F961

Landschützer, P., Gruber, N., & Bakker, D. C. (2016). Decadal variations and trends of the global ocean carbon sink. *Global Biogeochemical Cycles*, *30*(10), 1396–1417. https://doi.org/10.1002/2015GB005359

Landschützer, P., Gruber, N., Bakker, D. C., Schuster, U., Nakaoka, S., Payne, M. R., et al. (2013). A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences*, *10*(11), 7793–7815. https://doi.org/10.5194/bg-10-7793-2013

Landschützer, P., Gruber, N., Bakker, D. C., Stemmler, I., & Six, K. D. (2018). Strengthening seasonal marine CO_2 variations due to increasing atmospheric CO_2. *Nature Climate Change*, *8*(February), 146–150. https://doi.org/10.1038/s41558-017-0057-x

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C., van Heuven, S., et al. (2015). The reinvigoration of the Southern Ocean carbon sink. *Science*, *349*(6253), 1221–1224. https://doi.org/10.1126/science.aab2620

Laruelle, G. G., Landschützer, P., Gruber, N., Ti, J. L., Delille, B., Regnier, P. A., & Tison, J.-L. (2017). Global high-resolution monthly pCO_2 climatology for the coastal ocean derived from neural network interpolation. *Biogeosciences*, *14*(19), 4545–4561. https://doi.org/10.5194/bg-14-4545-2017

Legeckis, R. (1977). Long waves in the eastern equatorial Pacific ocean: A view from a geostationary satellite. *Science*, *197*(4309), 1179–1181. https://doi.org/10.1126/science.197.4309.1179

Leinweber, A., Gruber, N., Frenzel, H., Friederich, G. E., & Chavez, F. P. (2009). Diurnal carbon cycling in the surface ocean and lower atmosphere of Santa Monica Bay, California. *Geophysical Research Letters*, *36*(8). https://doi.org/10.1029/2008GL037018

Lenton, A., Matear, R. J., Keller, D. P., Scott, V., & Vaughan, N. E. (2018). Assessing carbon dioxide removal through global and regional ocean alkalinization under high and low emission pathways. *Earth System Dynamics*, *9*(2), 339–357. https://doi.org/10.5194/esd-9-339-2018

Lévy, M., Lengaigne, M., Bopp, L., Vincent, E. M., Madec, G., Ethé, C., et al. (2012). Contribution of tropical cyclones to the air-sea CO_2 flux: A global view. *Global Biogeochemical Cycles*, *26*(2). https://doi.org/10.1029/2011GB004145

Liang, J.-H., McWilliams, J. C., & Gruber, N. (2009). High-frequency response of the ocean to mountain gap winds in the northeastern tropical Pacific. *Journal of Geophysical Research*, *114*(C12). https://doi.org/10.1029/2009JC005370

Ma, D., Gregor, L., & Gruber, N. (2023). Four decades of trends and drivers of global surface ocean acidification. *Global Biogeochemical Cycles*, *37*(7). https://doi.org/10.1029/2023GB007765

Mayot, N., Le Quéré, C., Rödenbeck, C., Bernardello, R., Bopp, L., Djeutchouang, L. M., et al. (2023). Climate-driven variability of the Southern Oceanocean CO_2 sink. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *381*(2249). https://doi.org/10.1098/rsta.2022.0055

McKinley, G. A., Fay, A. R., Eddebbar, Y. A., Gloege, L., & Lovenduski, N. S. (2020). External forcing explains recent decadal variability of the ocean carbon sink. *AGU Advances*, *1*(2), e2019AV000149. https://doi.org/10.1029/2019AV000149

Merchant, C. J., Embury, O., Bulgin, C. E., Block, T., Corlett, G. K., Fiedler, E., et al. (2019). Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Scientific Data*, *6*(1), 223. https://doi.org/10.1038/s41597-019-0236-x

Merlivat, L., Gonzalez Davila, M., Caniaux, G., Boutin, J., & Reverdin, G. (2009). Mesoscale and diel to monthly variability of CO_2 and carbon fluxes at the ocean surface in the northeastern Atlantic. *Journal of Geophysical Research*, *114*(C3). https://doi.org/10.1029/2007JC004657

Mignot, A., von Schuckmann, K., Landschützer, P., Gasparin, F., van Gennip, S., Perruche, C., et al. (2022). Decrease in air-sea CO_2 fluxes caused by persistent marine heatwaves. *Nature Communications*, *13*(1), 4300. https://doi.org/10.1038/s41467-022-31983-0

Monteiro, P. M. S., Gregor, L., Lévy, M., Maenner, S., Sabine, C. L., & Swart, S. (2015). Intraseasonal variability linked to sampling alias in air-sea CO_2 fluxes in the Southern Ocean. *Geophysical Research Letters*, *42*(20), 8507–8514. https://doi.org/10.1002/2015GL066009

Müller, J. D., Gruber, N., Carter, B., Feely, R., Ishii, M., Lange, N., et al. (2023). Decadal trends in the oceanic storage of anthropogenic carbon from 1994 to 2014. *AGU Advances*, *4*(4), e2023AV000875. https://doi.org/10.1029/2023AV000875

Müller, J. D., Schneider, B., & Rehder, G. (2016). Long-term alkalinity trends in the Baltic Sea and their implications for CO_2-induced acidification. *Limnology & Oceanography*, *61*(6), 1984–2002. https://doi.org/10.1002/lno.10349

Murphy, P. P., Nojiri, Y., Harrison, D. E., & Larkin, N. K. (2001). Scales of spatial variability for surface ocean pCO_2 in the Gulf of Alaska and Bering Sea: Toward a sampling strategy. *Geophysical Research Letters*, *28*(6), 1047–1050. https://doi.org/10.1029/2000GL012375

Naegler, T. (2009). Reconciliation of excess 14C-constrained global CO_2 piston velocity estimates. *Tellus Series B Chemical and Physical Meteorology*, *61 B*(2), 372–384. https://doi.org/10.1111/j.1600-0889.2008.00408.x

Nicholson, S. A., Whitt, D. B., Fer, I., du Plessis, M. D., Lebéhot, A. D., Swart, S., et al. (2022). Storms drive outgassing of CO2 in the subpolar Southern Ocean. *Nature Communications*, *13*(1), 1–12. https://doi.org/10.1038/s41467-021-27780-w

Palter, J. B., Nickford, S., & Mu, L. (2023). Ocean carbon dioxide uptake in the tailpipe of industrialized continents. *Geophysical Research Letters*, *50*(21), e2023GL104822. https://doi.org/10.1029/2023GL104822

Pardo, P. C., Tilbrook, B., van Ooijen, E., Passmore, A., Neill, C., Jansen, P., et al. (2019). Surface ocean carbon dioxide variability in South Pacific boundary currents and Subantarctic waters. *Scientific Reports*, *9*(1), 7592. https://doi.org/10.1038/s41598-019-44109-2

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., et al. (2022). FourCastNet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv:2202.11214*. https://doi.org/10.48550/arXiv.2202.11214

Regnier, P. A., Resplandy, L., Najjar, R. G., & Ciais, P. (2022). The land-to-ocean loops of the global carbon cycle. *Nature*, *603*(7901), 401–410. https://doi.org/10.1038/s41586-021-04339-9

Resplandy, L., Hogikyan, A., Müller, J. D., Najjar, R. G., Bange, H. W., Bianchi, D., et al. (2024). A synthesis of global coastal ocean greenhouse gas fluxes. *Global Biogeochemical Cycles*, *38*(1), e2023GB007803. https://doi.org/10.1029/2023GB007803

Reul, N., Chapron, B., Grodsky, S. A., Guimbard, S., Kudryavtsev, V., Foltz, G. R., & Balaguru, K. (2021). Satellite observations of the sea surface salinity response to tropical cyclones. *Geophysical Research Letters*, *48*(1), e2020GL091478. https://doi.org/10.1029/2020GL091478

Ritter, R., Landschützer, P., Gruber, N., Fay, A. R., Iida, Y., Jones, S. D., et al. (2017). Observation-based trends of the southern Ocean Carbon sink. *Geophysical Research Letters*, *44*(24), 12339–12348. https://doi.org/10.1002/2017GL074837

Rödenbeck, C., Bakker, D. C., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S. D., et al. (2015). Data-based estimates of the ocean carbon sink variability - First results of the surface ocean pco2 mapping intercomparison (SOCOM). *Biogeosciences*, *12*(23), 7251–7278. https://doi.org/10.5194/bg-12-7251-2015

Rödenbeck, C., Bakker, D. C., Metzl, N., Olsen, A., Sabine, C. L., Cassar, N., et al. (2014). Interannual sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme. *Biogeosciences*, *11*(17), 4599–4613. https://doi.org/10.5194/bg-11-4599-2014

Rödenbeck, C., DeVries, T., Hauck, J., Le Quéré, C., & Keeling, R. F. (2022). Data-based estimates of interannual sea–air CO_2 flux variations 1957–2020 and their relation to environmental drivers. *Biogeosciences*, *19*(10), 2627–2652. https://doi.org/10.5194/bg-19-2627-2022

Rodgers, K. B., Schwinger, J., Fassbender, A. J., Landschützer, P., Yamaguchi, R., Frenzel, H., et al. (2023). Seasonal variability of the surface ocean carbon cycle: A synthesis. *Global Biogeochemical Cycles*, *37*(9), e2023GB007798. https://doi.org/10.1029/2023GB007798

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models (No. arXiv:2112.10752). *arXiv*. https://doi.org/10.48550/arXiv.2112.10752

Romero-Centeno, R., Zavala-Hidalgo, J., Gallegos, A., & O'Brien, J. J. (2003). Isthmus of Tehuantepec wind climatology and ENSO signal. *Journal of Climate*, *16*(15), 2628–2639. https://doi.org/10.1175/1520-0442(2003)016⟨2628:IOTWCA⟩2.0.CO;2

Roobaert, A., Regnier, P., Landschützer, P., & Laruelle, G. G. (2023). *Earth System Science Data Discussions*, 1–32. https://doi.org/10.5194/essd-2023-228

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., et al. (2004). The oceanic sink for anthropogenic CO2. *Science*, *305*(5682), 367–371. https://doi.org/10.1126/science.1097403

Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C., Pfeil, B., Olsen, A., et al. (2013). Surface Ocean CO2 Atlas (SOCAT) gridded data products. *Earth System Science Data*, *5*(1), 145–153. https://doi.org/10.5194/essd-5-145-2013

Sarmiento, J. L., & Gruber, N. (2006). *Ocean biogeochemical dynamics*. Princeton University Press.

Sathyendranath, S., Jackson, T., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., et al. (2023). ESA Ocean colour climate change initiative (Ocean_Colour_cci): Version 6.0, 4km resolution data. [Dataset]. *NERC EDS Centre for Environmental Data Analysis*. https://doi.org/10.5285/5011D22AAE5A4671B0CBC7D05C56C4F0

Sharp, J. D., Fassbender, A. J., Carter, B. R., Lavin, P. D., & Sutton, A. J. (2022). A monthly surface pCO_2 product for the California current large marine ecosystem. *Earth System Science Data*, *14*(4), 2081–2108. https://doi.org/10.5194/essd-14-2081-2022

Shutler, J. D., Gruber, N., Findlay, H. S., Land, P. E., Gregor, L., Holding, T., et al. (2024). The increasing importance of satellite observations to assess the ocean carbon sink and ocean acidification. *Earth-Science Reviews*, *250*, 104682. https://doi.org/10.1016/j.earscirev.2024.104682

Shutler, J. D., Wanninkhof, R. H., Nightingale, P. D., Woolf, D. K., Bakker, D. C., Watson, A. J., et al. (2020). Satellites will address critical science priorities for quantifying ocean carbon. *Frontiers in Ecology and the Environment*, *18*(1), 27–35. https://doi.org/10.1002/fee.2129

Song, Z., Yu, S., Bai, Y., Guo, X., He, X., Zhai, W., & Dai, M. (2023). Construction of a high spatiotemporal resolution dataset of satellite-derived pCO_2 and air–sea CO_2 flux in the South China sea (2003–2019). *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–15. https://doi.org/10.1109/TGRS.2023.3306389

Sutton, A. J., Feely, R. A., Maenner-Jones, S., Musielwicz, S., Osborne, J., Dietrich, C., et al. (2019). Autonomous seawater pCO 2 and pH time series from 40 surface buoys and the emergence of anthropogenic trends. *Earth System Science Data*, *11*(1), 421–439. https://doi.org/10.5194/essd-11-421-2019

Sutton, A. J., Sabine, C. L., Maenner-Jones, S., Lawrence-Slavas, N., Meinig, C., Feely, R. A., et al. (2014). A high-frequency atmospheric and seawater pCO_2 data set from 14 open-ocean sites using a moored autonomous system. *Earth System Science Data*, *6*(2), 353–366. https://doi.org/10.5194/essd-6-353-2014

Sutton, A. J., Wanninkhof, R. H., Sabine, C. L., Feely, R. A., Cronin, M. F., & Weller, R. A. (2017). Variability and trends in surface seawater pCO_2 and CO_2 flux in the Pacific Ocean. *Geophysical Research Letters*, *44*(11), 5627–5636. https://doi.org/10.1002/2017GL073814

Sutton, A. J., Williams, N. L., & Tilbrook, B. (2021). Constraining Southern Oceanocean CO_2 flux uncertainty using uncrewed surface vehicle observations. *Geophysical Research Letters*, *48*(3), 1–9. https://doi.org/10.1029/2020GL091748

Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G. A., Sarmiento, J. L., & Wanninkhof, R. H. (2007). Constraining global air-sea gas exchange for CO_2 with recent bomb ^{14}C measurements. *Global Biogeochemical Cycles*, *21*(2), 1–10. https://doi.org/10.1029/2006GB002784

Taburet, G., Sanchez-Roman, A., Ballarotta, M., Pujol, M.-I., Legeais, J.-F., Fournier, F., et al. (2019). Duacs DT2018: 25 years of reprocessed sea level altimetry products. *Ocean Science*, *15*(5), 1207–1224. https://doi.org/10.5194/os-15-1207-2019

Tans, P. P., & Keeling, R. F. (2023). Global monitoring laboratory - carbon cycle greenhouse gases [Dataset]. Retrieved from https://gml.noaa.gov/ccgg/trends/

Telszewski, M., Chazottes, A., Schuster, U., Watson, A. J., Moulin, C., Bakker, D. C. E., et al. (2009). Estimating the monthly pCO_2 distribution in the North Atlantic using a self-organizing neural network. *Biogeosciences*, *6*(8), 1405–1421. https://doi.org/10.5194/bg-6-1405-2009

Toolsee, T., Nicholson, S.-A., & Monteiro, P. M. S. (2024). Storm-driven pCO_2 feedback weakens the response of air-sea CO_2 fluxes in the sub-Antarctic Southern Ocean. *Geophysical Research Letters*, *51*(9), e2023GL107804. https://doi.org/10.1029/2023GL107804

Torres, O., Kwiatkowski, L., Sutton, A. J., Dorey, N., & Orr, J. C. (2021). Characterizing mean and extreme diurnal variability of Oceanocean CO_2 system variables across marine environments. *Geophysical Research Letters*, *48*(5), 1–12. https://doi.org/10.1029/2020GL090228

Turi, G., Lachkar, Z., & Gruber, N. (2014). Spatiotemporal variability and drivers of pCO_2 and air–sea CO_2 fluxes in the California Current System: An eddy-resolving modeling study. *Biogeosciences*, *11*(3), 671–690. https://doi.org/10.5194/bg-11-671-2014

Wanninkhof, R., Pickers, P. A., Omar, A. M., Sutton, A., Murata, A., Olsen, A., et al. (2019). A Surface Ocean CO2 reference network, SOCONET and associated marine boundary layer CO_2 measurements. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00400

Wanninkhof, R. H. (2014). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and Oceanography: Methods*, *12*(6), 351–362. https://doi.org/10.4319/lom.2014.12.351

Weiss, R. (1974). Carbon dioxide in water and seawater: The solubility of a non-ideal gas. *Marine Chemistry*, *2*(3), 203–215. https://doi.org/10.1016/0304-4203(74)90015-2

Weiss, R., & Price, B. (1980). Nitrous oxide solubility in water and seawater. *Marine Chemistry*, *8*(4), 347–359. https://doi.org/10.1016/0304-4203(80)90024-9

Whitt, D. B., Nicholson, S. A., & Carranza, M. M. (2019). Global impacts of subseasonal (<60 Day) wind variability on ocean surface stress, buoyancy flux, and mixed layer depth. *Journal of Geophysical Research: Oceans*, *124*(12), 8798–8831. https://doi.org/10.1029/2019JC015166

Woolf, D. K., Land, P. E., Shutler, J. D., Goddijn-Murphy, L., & Donlon, C. J. (2016). On the calculation of air-sea fluxes of CO_2 in the presence of temperature and salinity gradients. *Journal of Geophysical Research: Oceans*, *121*(2), 1229–1248. https://doi.org/10.1002/2015JC011427

Ye, H., Morozov, E., Tang, D., Wang, S., Liu, Y., Li, Y., & Tang, S. (2020). Variation of pCO_2 concentrations induced by tropical cyclones "wind-pump" in the middlelatitude surface oceans: A comparative study. *PLoS One*, *15*(3), e0226189. https://doi.org/10.1371/journal.pone.0226189

Yu, P., Wang, Z. A., Churchill, J., Zheng, M., Pan, J., Bai, Y., & Liang, C. (2020). Effects of typhoons on surface seawater pCO_2 and air-sea CO_2 fluxes in the northern south China sea. *Journal of Geophysical Research: Oceans*, *125*(8). https://doi.org/10.1029/2020JC016258